



Article

Design for an Art Therapy Robot: An Explorative Review of the Theoretical Foundations for Engaging in Emotional and Creative Painting with a Robot

Martin Daniel Cooney * and Maria Luiza Recena Menezes

Intelligent Systems and Digital Design Department, Halmstad University, P.O. Box 823, Kristian IV:s väg 3, 30118 Halmstad, Sweden; maria.menezes@hh.se

* Correspondence: martin.daniel.cooney@gmail.com or marcoo@hh.se

Received: 18 July 2018; Accepted: 28 August 2018; Published: 3 September 2018



Abstract: Social robots are being designed to help support people's well-being in domestic and public environments. To address increasing incidences of psychological and emotional difficulties such as loneliness, and a shortage of human healthcare workers, we believe that robots will also play a useful role in engaging with people in therapy, on an emotional and creative level, e.g., in music, drama, playing, and art therapy. Here, we focus on the latter case, on an autonomous robot capable of painting with a person. A challenge is that the theoretical foundations are highly complex; we are only just beginning ourselves to understand emotions and creativity in human science, which have been described as highly important challenges in artificial intelligence. To gain insight, we review some of the literature on robots used for therapy and art, potential strategies for interacting, and mechanisms for expressing emotions and creativity. In doing so, we also suggest the usefulness of the responsive art approach as a starting point for art therapy robots, describe a perceived gap between our understanding of emotions in human science and what is currently typically being addressed in engineering studies, and identify some potential ethical pitfalls and solutions for avoiding them. Based on our arguments, we propose a design for an art therapy robot, also discussing a simplified prototype implementation, toward informing future work in the area.

Keywords: social robots; art therapy; emotions; creativity; art robots; therapy robots

1. Introduction

The current article presents a basic design for an art therapy robot. Specifically, we derive some guidelines from the literature in exploring *how to visually express emotions in a creative context* of painting with a robot, as a step toward realizing a therapy robot which could positively support people's well-being by engaging with people on an emotional and creative level; additionally, we provide an implementation example from some of our work in recent years, illustrated in Figure 1.

1.1. Terms

"Art therapy" is a therapeutic process involving art-making: a patient expresses emotions through creating art, which also serves as a bridge between the patient and a therapist [1]. "Therapy" here encompasses notions of care, healing, and providing attention [2] and is intended to mitigate problems and facilitate positive health and well-being, where "well-being" here is used synonymously with "happiness", "quality of life", and "life-satisfaction" [3]. "Art", which has been described as washing "away from the soul the dust of everyday life" [4], can include painting, drawing, photography, collage, and sculpture, and be abstract or symbolic, where we use the term "symbol" here in its general sense as some meaningful stimulus pattern, thereby comprising icons and indices. "Emotion" in humans

refers to a complex psycho-physical phenomenon at the heart of art therapy [5,6], involving subjective feelings, somatic symptoms such as elevated heart rate, affect displays such as smiling, and cognitive appraisals [7]; we suggest that some interesting properties of emotions to consider for art therapy include co-occurrence, referents, timing, and polysemy (that emotions often co-occur, that emotions usually refer to some referent which can be someone or something, that emotions play out over time, and that emotion displays can express different emotions), as described in Appendix A. “Creativity” is another core aspect of art therapy, which is “a way of doing things” [8], characterized by some relative novelty, as described in Appendix B. Additionally, for “robot” here, we primarily consider (semi-)autonomous machines, comprising computers, sensors and actuators, with some degree of human-like intelligence and capabilities that can be used for therapy.



Figure 1. The basic concept of the current work is that we imagine robots can play a useful role in engaging with people emotionally and creatively in art therapy. The photo shows a robot painting with a person, which bases its artwork on the person’s emotions inferred using a Brain–Machine Interface. Used here with permission.

1.2. Motivation

There is a large and rising need to help people mentally: for example, loneliness, which has been tied also to dementia [9], has been described as a “rising epidemic” that can be a higher mortality risk than moderate daily smoking or obesity when prolonged, and is estimated to cost hundreds of billions of dollars yearly in the US alone [10]. One useful approach for mentally helping patients is art therapy, which has demonstrated some good effects with a wide range of groups, as follows. Some positive subjective effects of art therapy include improved self-awareness, self-image, relaxation, and social interactions; objectively measured outcomes include improved clinical outcomes, better vital signs, reduced cortisol, better sleep, shorter stays, faster discharges, and higher pain tolerance [11]. Furthermore, effects can persist; for example, “a subtle, more pronounced and durable positive effect across time” of art therapy was noted compared to an alternative of recreational activity group work in dementia patients [12]. It has even been suggested that art therapy potentially has some “preventive, diagnostic, therapeutic and rehabilitative benefits that other forms of therapy

cannot provide” as art offers another rich way for people to explore themselves other than through the spoken word or playing [1]. (We note that this by no means implies that art therapy is better than other forms of therapy, which can also have unique and powerful benefits; for instance, in robotics, some studies using Paro and KiliRo have reported success in reducing stress and psychologically helping elderly and children with autism [13–15].) Art therapy has been used with both adults and children for various conditions, including dementia, autism, depression, trauma (post-traumatic stress disorder, sexual abuse, and traumatic brain injury), AD/HD, schizophrenia, bipolar disorder, borderline personality disorder, AIDS, asthma, burns, cancer, chemical dependency, hemodialysis, sickle cell disease, and tuberculosis [6,16–19].

To provide art therapy, we believe that a painting robot could potentially be useful: for example, helping people who cannot currently receive care due to a shortage of human resources [20] as robots can be manufactured as needed and programmed knowledge can be shared; saving some time for human therapists, such as for traveling to, attending, and documenting sessions; leveraging abilities not normally available to humans, such as inferring emotions from brainwaves; being available at possibly any time as robots do not need to sleep; and facilitating self-exploration without requiring people to express vulnerable thoughts to another human. In practice, some benefits of therapy robots have already been described in other contexts, such as decreasing loneliness in persons with dementia with pet-like robots, Paro and AIBO [21,22]. Furthermore, in comparison with other technologies such as virtual agents, robots have been reported to elicit stronger perceived emotions, presence, motivation, and engagement [23–26], and could perform more behaviors, such as seeking out a person to interact, working in 3D, and conducting other useful tasks in a home such as cleaning and fetching items. Additionally, painting is a typical form of art which engrosses multiple senses—not only sight and touch but to a lesser degree also smell and sound, allows for high creativity as objects do not have to be seen to be painted; does not require knowledge of computers; and leaves a tangible, persisting artifact which can be a reminder of achievement. Furthermore, affective computing and artificial creativity relate to the strategic area of artificial intelligence, where multi-billion dollar plans are being proposed and it has been claimed that “whoever becomes the leader in this sphere will become the ruler of the world” [27].

Thus, the current work seeks to present a design for a therapy robot, which could positively support people’s well-being by engaging with people on an emotional and creative level. A challenge is that the theoretical foundations are highly complex: robot art therapy requires gathering knowledge from various fields in human science and engineering; and, as noted, we are only just beginning ourselves to understand emotions and creativity in human science, which have been described as important topics, final frontiers and ultimate challenges in artificial intelligence [28–30].

To gain insight, our approach is to explore the theoretical foundations by reviewing some of the extensive literature, identifying issues we feel to be important, proposing prescriptions, and clarifying possibilities, in Section 2. (We note that, in Section 2, the current article does not seek to provide a definitive review of a single, existing, mature field, but rather seeks to explore what valuable information can be drawn together from some online sources in multiple fields to design a new interaction, involving robot art therapy. In terms of the PICO paradigm used in medical studies, we include references to work conducted with various participant groups, as we believe that various groups could benefit from robot art therapy, and that interaction by a robot with various groups will be important to avoid stigmatization. Interventions of interest relate both to human art therapy and to general robot interaction strategies, whereas outcomes of interest relate to facilitation of well-being and positive emotions. Likewise, we do not place restrictions on year of publication; however, only publications in English language are considered.) Based on our arguments, we propose a design for an art therapy robot, also discussing a simplified prototype implementation in Section 3 (some code has been made available on GitHub (<https://github.com/martincooney>), and some videos on YouTube (<https://youtu.be/9d96MrSCjx8>, and other footage of our robot painting at https://www.youtube.com/channel/UCVDPmL7NkC5Mn_A1wU0RPKg)). Section 4 provides some

extra discussion, toward informing future work in the area. Additionally, we note that a portion of the current article extends and directly builds upon some of our recent work on identifying some general potential dangers in affective computing [31], and implementing emotional art-making behaviors [32].

2. Theoretical Foundations

To propose a design for an art therapy robot, focusing on emotions and creativity, we believe the first important step is to consider the theoretical foundations. Here, we review some related work, identifying issues we feel to be important, and proposing prescriptions. Related work concerns robots used for therapy and art, strategies for interacting in art therapy, and mechanisms for expressing emotions and creativity.

2.1. Therapy Robots

Numerous kinds of therapy exist—e.g., physical, play, dance, music, drama, writing, gardening, and art—some of which are already being conducted with various robots.

In particular, physical therapy robots have a long history. Rudimentary prosthetics and sensors might have been used since the times of Ancient Egypt [33]; more recently, exoskeletons were designed as tools to help people walk in the 1800s, and to carry heavy loads since the 1960s [34]. Such robots have also begun to be used more recently in other roles, such as “socially assistive” robots working as exercise coaches, to get elderly to move their arms [35]. Another robot was built to act as an “opponent”; elderly chase the robot, whose behavior adapts based on perceived skill levels [36]. Furthermore, a robot was constructed to be an exercise partner; elderly walk with the robot between fixed stations by clicking a button to control the robot to move [37].

In addition, many playful robots have been built which are intended to function like pets, toys, tools, or play partners. Playful pet robots, such as the seal-like robot Paro and dog-like robot AIBO, which were introduced in the 1990s, and more recently the teddy bear-like robot CUDDler, have been used in place of animals in “Robot Assisted Therapy” (RAT) [21,22,38]. Robots have also been built as toys for children to program, such as PETS, which can tell stories [39]. Playful parrot robots, RoboParrot and KiliRo, have been used with autistic children, as tools for screening and to achieve enjoyable learning [13,40]. Furthermore, some robots have been built as play partners, such as Maggie, which can play various games including Tic-tac-toe, Hangman, Twenty Questions, Peekaboo, and “Animal Quiz” [41], or Probo, Zeno, Rovio, and Nao, which have also been used to play games with autistic children, involving identifying emotions [42,43] and imitation [44,45].

In addition, some work has explored designs for robots which could be used for dance or music therapy. For example, Kosuge et al. created a robot which can dance with people as a dance partner [46], and Kozima et al. built Keepon to investigate the usefulness of rhythm synchronicity, which could be used to enable therapeutic dance interactions [47]. In addition, an interaction with a music therapist robot was proposed, in which people guess from hearing a few notes which song a robot is playing [48]; robots which can interact musically with humans have also been built [49], which could also be used for therapy.

Additionally some robots can engage in more than one kind of therapy. For example, Martín et al. conducted tests with dementia patients using a Nao Robot designed for four types of sessions: physiotherapy, play with storytelling, music, and language (asking about numbers and days of week to stimulate thinking and recall) [50]. In addition, various other general healthcare robots have been built, which can carry, monitor, and fetch items for humans [51–53]. Moreover, recently a need to develop more autonomous therapy robots has been described by researchers in the DREAM project, using the term “Robot enhanced therapy” (RET); their goal is to use robots as a tool to help children with autism to improve social skills comprising turn-taking, joint attention, imitation, in conjunction with Applied Behavioral Analysis (ABA), a learning theory based on behavioral repetition and cognitive association [54]. Art therapy could also potentially be structured to improve such social skills. Another interesting proposal has been made that robots could conduct “alloparenting”, acting as

surrogate guardians for human children, which we believe also relates to therapy, as parents are also required to support the emotional and mental well-being of children; this work highlights the importance of joint attention and empathy, as well as the promotion of autonomy through principles of beneficence, non-maleficence, equal provision of resources, and respect for a person’s dignity [55]. In summary, therapy robotics appears to be an active area of research in which various work is being done, with both children and elderly, and also persons with physical disabilities, dementia, and autism [56,57].

Some gaps seem to exist for therapy regarding drama, writing, gardening, and art therapy. For example, a robot could read lines from a play with a human, write a diary together with a human (e.g., filling in things which a patient has forgotten to mention, or asking for further details), garden (e.g., planting or harvesting with a human), or make art with a human. We focus here on the latter possibility; to our knowledge, previous work has not proposed a design for an art therapy robot.

2.2. Robot Art

Various machines have been built to create visual art, under names such as computer art, generative art, and electronic art: typically computers are used to generate images, which can be translated into physical art by printers and robot embodiments. Some first artistic images on computers were reportedly created for fun by IBM employees in the 1950s [58]. In the 1970s, an artist named Cohen started producing paintings generated by a computer program, AARON, which were executed by plotters and printers [59]. In the 1990s, a drawing robot called ISAC was also designed to mimic an artist’s movements [60]. In the last decades, various detailed prescriptions were made, such as how a robot can grasp a paint brush in its fingers, and paint, detecting the brush tip [61], and numerous robots have been designed as tools to draw some existing scene in an aesthetically appealing way. Recently, every year various such robots take part in the international robot art competition (robotart.org), illustrated in Figure 2, and there is also a “DeepArt” competition for style transfer artwork (at the Annual Conference on Neural Information Processing Systems (<https://deepart.io/nips/submissions/>)). Such systems usually seek to create art that is aesthetically pleasing by taking a model or an image, possibly altered via style transference, as input, and using reinforcement learning to iteratively refine artwork, or mimicking an artist’s movements directly. However, for robot art therapy, we expect that there might not be a photo to compare to, work might be only partially completed, and robots should move not only based on an artist’s motions.

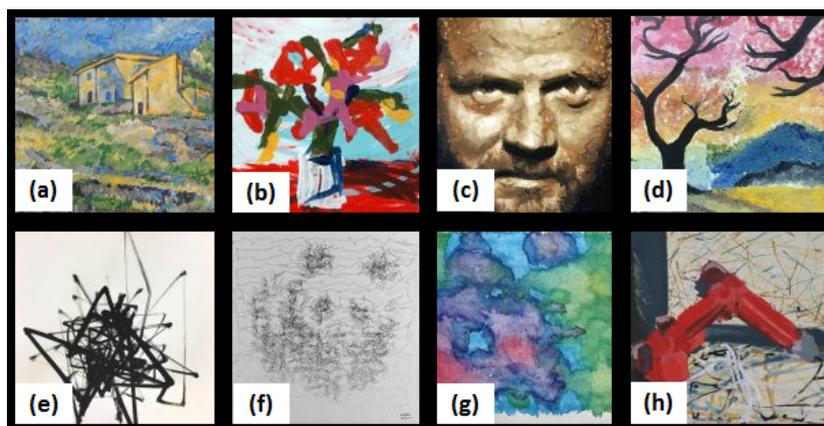


Figure 2. Examples of paintings from the robot art competition: (a) Cezanne’s Houses at L’Estaque by CloudPainter; (b) Red Flowers (Floral no. 1) by Joanne Hastie; (c) Man by PIX 18 at Columbia University; (d) Full Bloom of Sakura by CMIT ReART at Kasetsart University; (e) Scribbles by CARP at Worcester Polytechnic Institute; (f) WWF by JACKbDU at New York University Shanghai; (g) Perlin Noise Field by Late Night Projects; and (h) Homage To Jackson Pollock by e-David at University of Konstanz.

Some interactive art systems have also been created which are not controlled by an artist's movements. For example, one humanoid robot asked people if they would like to be drawn, and then drew detected faces step-by-step, from rough contours to more detailed lines [62]; this interaction resembles our vision for how a robot can interact with a human not as a tool but as an interaction partner, while co-creating art. In addition, recently, in the field of human-computer interaction, some interactive systems have been developed. Lee et al. described an interface for assisted drawing using a tablet, where the system used shadows to provide suggestions on how to complete sketches [63]; likewise, Ha et al. made available online a tool called "sketch-rnn" which can also generate such suggestions, using a Sequence-to-Sequence Variational Autoencoder [64]. We note that, in practice, suggestions are often drawn from where a person has just finished drawing, which could interfere in a physical system: a robot should not push aside a person's hand to start painting in the same spot.

In addition, we suggest that, to go further, a robot could recognize which "rules" a person is following when painting (which might relate to speed or placement within a composition), predict intentions and next moves, track its performance, and plan how to behave next. Regarding the first problem of aligning a robot's work with a human's, previous work has demonstrated classification of different styles of painting such as baroque or impressionist, which could possibly be one parameter to consider [65]. Furthermore, Crick et al. described a camera-equipped robot which could learn the rules of a game (i.e., the intentions of other robots) via a dynamics-based causal model, and then participate using its motions to resolve ambiguities and further improve its understanding [66]. For prediction, Mutlu et al. reported benefits of having a robot predict what a human will do next, using a person's gaze, in a scenario about a robot handing objects to a human [67]. Montebelli et al. also reported possibilities for recognizing intentions through motion patterns, although highlighting individualistic differences [68]. Furthermore, work has been done on looking toward a human for feedback (e.g., Breazeal's Leonardo [69]), and use of smiling for reinforcement [70], although in these works the referent has been clear, and smiling by itself can result from various reasons. Some inspiration can be drawn from such work, on how an art robot could extend a human's painting, and engage in harmonious co-creation, but the focus has been largely not been on therapy, emotions, and well-being.

2.3. Interaction Strategies for Art Therapy

As described in the previous section, various robots have been proposed for physical exercise and play, and art robots have been built to create artwork which is optimally pleasing to the eye, but the literature does not clearly indicate how a robot can interact to support a human seeking to engage in art therapy. Even turning for inspiration to human science to examine how art therapy is typically conducted, we encounter the challenge that there is no one "accepted" way to conduct art therapy; rather numerous approaches exist [71], some of which are briefly described in Appendix C. Below, we first clarify a proposal for a basic scenario, based on our own ideas. Then, we describe some ideas from human science, from some explicit guidelines, to some hints implied by underlying mechanisms in art therapy, before considering some general guidelines.

2.3.1. Humanistic Art Therapy: Robot as Partner

In the current article, we argue that some approaches might not be appropriate for robots, due to the risk of harm if a robot presents a mistaken diagnosis. First, mistakes could be common in first robot prototypes; art therapy requires a high degree of understanding of humans, which would be currently difficult to expect from a robot given the emerging nature of this field. Second, there seems to be a potentially dangerous power imbalance in the case of a "normal" therapist looking over the shoulder of an "abnormal" patient, judging, in that decisions could be made based on some diagnosis which could adversely affect a person. Some similar arguments have been made in human science: for example, it has been proposed that some approaches could be oppressive if there is some disagreement in values between therapist and patient [72,73]. This line of thought was also echoed in the area of human-robot interaction, for therapy in general, by Ziemke and colleagues, who questioned the assumption that

the therapist is an all-knowing expert who can deduce the truth [54], and by Tapus et al. who also advocated that therapy robots should be hands-off [74]. Furthermore Kahn et al. stated that the important target is how to design a scenario in which people will interact with robots as partners in a joint creative enterprise [75]. Thus, we envision, from a humanistic perspective, that the first fundamental scenario to target for art therapy robots should be one in which the person is at the center of the interaction, and the robot does not diagnose and judge, but rather tries to play a supporting role.

Following this basic conceptualization that a robot can act as a partner in art rather than a judge, we propose a basic scenario, following the Five W strategy (why, who, what, where, when) to address some salient questions [76]. Since “why” has already been discussed, we turn first to the question of who should paint. Art therapy can be conducted with various numbers of people and robots. A benefit of interacting with a single person, rather than a group, is that we expect the person will be better able to perceive full attention from the robot, due to, e.g., the “Socratic bottleneck” where in a group only one person can speak at once [77]. (If we assume the art therapy robot uses a familiar interface for communicating—humanoid or possibly animal-like—then, similar humans or animals, it can show attention through cues such as gaze, body pose, speech, location, and motions (e.g., art-making), but, because humans and animals tend to only have one head, one body, and a few arms, and take turns speaking, such a robot cannot simultaneously directly look at and listen to multiple people at once. It would have to look from one to another, listening to one person’s response then another’s, similar to a human, providing an impression of giving less than its full attention to each person. In other words, as an example, any time one person is speaking is a time when another person will appear to not be receiving full attention. This effect is known in pedagogy, and a reason for sometimes breaking students apart into small groups, so that they can more actively interact [77]).

Group therapy can also result in a sharing of strongly emotional experiences which might otherwise not be possible, and improve relationships [1]; however, this is a more complex case, introducing new dynamics, such as the relationships between people and how choices will be regulated based on who is present. It could also be possible to allocate multiple robots per person; for example, one robot could steady the arm of a person with Parkinson’s disease, while another robot could paint beside them, providing company. However, we believe this case is also more complex. Therefore, we suggest that the more basic dyadic case is useful for initial explorations.

For such a dyadic case, we propose that three basic cases can be described, in which the robot’s involvement in painting is 0% (only the person paints), somewhere in between 0% and 100%, such as 50% (both the person and robot paint), or 100% (only the robot paints). To determine which a person prefers, the robot can ask at the beginning of the interaction. In the first case, the person might wish to achieve independently without relying on others and feel ownership over the art. In the second case, the person might value the enjoyment from social interaction or expect a nicer result if co-creating. In the third case, the person might be incapable of physically participating, or merely prefer to passively observe. In the first and second cases, the robot can seek to infer a person’s emotions from the art they make, while, in the third case, the robot can seek to recognize how the person is feeling directly. In all cases, the robot can attempt to make some basic conversation: In the first and second cases, the robot can also ask the person about what they are painting, whereas, in the third case, the robot will simply comment on what it is painting. From the perspective of simplicity, the first case is arguably the simplest; however, the second and third cases we feel are most in line with our vision of the robot as a partner, and fundamental from the perspective of exploring how a robot can interact in an emotional and creative manner.

Regarding what a person should paint, we suggest giving the person themselves the dignity to freely select what to paint. Where a more formal procedure is desired, objective tests could be adapted, such as the Diagnostic Drawing Series (DDS) which assesses colors, lines, and composition (e.g., placement and integration), in asking a person to make a picture, make a picture of a tree, and make a picture of how they feel using lines, shapes, and colors [78]. However, we note that this test would not be possible to directly apply for the painting scenario considered here, as it is

designed specifically for pastels; also, requirements on time, number of artworks to produce, and subject matter, might not be desired in various cases: for example, if the artwork generated can also be intended to be aesthetic or displayed somewhere, or if a person requires more time, or does not have the physical strength to produce many drawings. We suggest that a robot can in general seek to track such objective information over time to roughly gauge a person's state and the effectiveness of sessions, even with freely chosen art. In addition, for individuals who cannot decide what to paint, a robot could suggest some topic; for example, self-portraiture is a tool used by art therapists as a way to promote self-reflectance and self-acceptance [79].

Another question which should be addressed is where the robot should paint. If a robot draws on the same substrate or canvas, the interaction could be felt to be more social and intimate; that is, there could be a stronger effect of social facilitation [80]. Furthermore, the robot could help the human to paint better, especially for individuals with restricted mobility, e.g., adding details which might require high dexterity, knowledge, or technical ability. Conversely, allowing a person to complete their own painting could result in an increased perception of accomplishment, and ownership, resulting from high perceived involvement, personalization, expression of territoriality, and control [81]. (Feeling ownership can also enhance memory through the so-called self-referential effect [82], which could be useful for dementia patients.) Such a case might also be simpler for initial investigations, as a robot does not need to track where a person is painting and avoid colliding with them. We argue that both cases can have benefits, so an art therapy robot should be capable of engaging in a range of art-making behaviors, from facilitation to completing a painting by itself. We also propose that the robot first ask the human how they would like to paint, and, if the same substrate is used, that the robot should allow the human as much as possible to play an important role in making the artwork.

In addition, a decision should be made on when art therapy should take place. Art therapy is tailored for the individual both in terms of number of sessions and structure of sessions, as follows [6]. Single sessions are possible, although therapy tends to last over several weeks to a year. Session can be approximately an hour each. A common structure for each session, in line with models such as the Creative Axis Model, is to have some warm-up activity, a main activity, and reflection at the end. Similarly, initial sessions can focus on goal-setting, identifying problems, and accustomization; middle sessions can involve working on central themes, then more challenging and complex themes, distancing from problems, investigating solutions and answering questions; and, the end can focus on review, next steps, and closure. In addition, art therapists can make responsive art before, during, or after sessions. We believe the most practical starting point for research in art therapy is focusing on a single session first. If the robot does not paint during the same session, timing and memory can be factors which should also be considered, which might be important, especially for persons with dementia; therefore, we also suggest that initial work focuses on the basic case in which the robot makes art at the same time as the person, in close social proximity.

Within this basic scenario, summarized in Table 1, we next consider some guidelines about how to conduct the desired form of humanistic art therapy. Phillips recommended: (1) investigating the meaning of the art with the person; (2) accepting (and at times encouraging) the communication of strong emotions, including negative ones; (3) praising creativity and skill, even for negative depictions; and (4) suggesting alternatives for disturbing negative content to express feelings [83]. Additionally, the importance of keeping track of the timing and progression in the artworks over time was mentioned, as well as the importance of understanding popular culture, clinical, and social context for seeking to find meaning in a person's art, although the latter could be difficult for a robot. These guidelines can be followed both via verbal interaction and via the therapist's own artwork, in a process which has been described as *responsive art*, or *visual feedback*. Phillips noted that often visual feedback was more successful than a verbal response in investigating meaning, in line with our idea that a robot can paint with a person as a companion.

Table 1. Proposal for a basic starting scenario for art therapy robot.

Question	Proposal
Who	One robot engages with one human (dyadic)
What	In nondirective painting
Where	Using different canvases
When	During a single session (1 h; comprising a warm-up, a main activity, and reflection)

Such prescriptions are in line with our idea of the robot acting as a companion, rather than a judge, but they do not clearly state how a robot can decide what to draw. For example, if a person is expressing a negative emotion such as sadness, what should a robot do? The literature presents some evidence supporting two potentially opposing premises: that a robot could try to *match* a human's negative emotions, and that it could seek to *distract* with a positive emotional display. In human-robot interaction, Goetz et al. compared a robot which is always positive to a robot which matched its mood to the context, finding the latter was more liked [84]; Tapus and colleagues also found that people preferred a robot to match its behaviors to a user's personality [85]. A benefit of such an approach could be that the person does not feel that they always have to be positive and suppress their emotions, which has been reported to be an ineffective emotion regulation strategy with some negative consequences [86]. Furthermore, a tendency to like people with similar attitudes has been described [87], implying that such convergence of emotion displays could engender liking. (We note that this concept of displaying similar emotions does not imply any suggestion that a robot's appearance should or should not resemble a human's, which is debated: According to Mori's Uncanny Valley hypothesis, a near-human appearance with slight imperfections could potentially trigger negative impressions; conversely, it has been argued that negative impressions can be avoided using an attractive design as not all imperfections elicit the same responses [88], and also that humans can quickly become accustomed and extend their expectations for appearance [89].) In addition, humor could also be perceived in causing a robot to behave in a sad manner, like seeing a pet's concern over its owner; negative emotions can also be valuable [90], and Philipp's second guideline above regarding accepting others' emotions could also be interpreted as suggesting matching.

A caveat is that the studies above were not conducted in the context of art therapy and it was not determined if people felt better as a result. In art therapy, Drake and Winner found that distraction had a better effect on a person's mood than venting (i.e., expressing positive, rather than negative emotions, which might help to avoid negative rumination) [91]. We believe various positive effects could ensue from a robot's positive emotional behavior, such as happiness through emotional contagion; a person could feel safe if a robot never displays negativity; the robot could seem to have its own goals and not be merely reactive; and, Philipp's third and fourth guidelines could be interpreted as suggesting distraction, by praising and suggesting alternatives.

However, Drake and Winner's study was not conducted with robots. Moreover, if a robot always acts in the same positive manner, ignoring the person's emotions, it could appear boring [92], or insincere [93]. Positive behaviors such as laughter can be irritating when perceived to stem from *schadenfreude* [94]. Moreover, if a robot adopts a purely positive stance about everything, and does not acknowledge any negative aspects, a person might feel the need to express the negative side themselves, in line with the concepts of the devil's advocate and reverse psychology, which might result in undesired negative rumination. Furthermore, the robot could be interpreted as implying that, if a person is not positive like it, they must have a problem, which could produce negative feelings.

We argue that the results of these studies do not necessarily contradict: for example, a robot in an interaction can do both, sometimes matching, thereby showing empathy and gaining trust, and sometimes distracting, thereby helping the user to feel better. Conversely, the robot could do both simultaneously, expressing a mixture of negative and positive emotions: for example, drawing a sad scene with a positive rainbow. Furthermore, there is not necessarily a conflict or disruption of contingency in responding positively to a negative emotional display. A study conducted in the

context of analyzing conversations on Twitter suggested that *the form in which positive emotions are shown is important*, finding that sympathy, greetings, and recommendations in particular exerted a strong positive influence on others' emotions [95]. Conversely, worrying, teasing, and complaining often caused others to feel negative emotions. In terms of basic Ekman emotion categories, worrying and complaint can be associated with fear and anger, respectively (high arousal, low valence), whereas sympathy is displayed through sadness (low arousal, low valence) toward a person's unfortunate situation, or affection (high valence) toward a person. Greetings and recommendations are typically happy (mid to high valence). Teasing is typically insincere, and can be insincere anger, or insincere joy. Additionally, the study also found that people usually expressed the same emotion or a positive emotion, in line with our idea that both matching and distraction are useful behaviors to employ. We note that there are also theories in human science such as communication accommodation theory (CAT) which could offer additional insight into when to match or distract (converge or diverge); however, this theory, in line with social identity theory (SIT), focuses on explaining human motivations for social approval, communication efficiency, and social identity, rather than on how a robot could engender positive affective changes [96].

We believe that a key concept underlying matching is empathy, which describes the capability to perceive, understand, and share another person's emotions [74,97]. From this perspective, we believe that matching is not merely mimicry in showing similar emotions, but rather there is also a cognitive aspect involving perspective taking, which is required to deal with complex emotions involving mixed emotions and referents. More specifically, when matching, emotions to display can be influenced by the type and degree of emotions perceived; the robot's personality; and the perceived importance for the robot to act based on the closeness of the relationship between robot and person, as well as the appraised context (e.g., for humans a bystander effect is observed in which helping behaviors are inhibited when many people are present) [98]. Based on such theory, a rich computational model of empathy has been built, which we believe will enable matching behavior in art therapy robots [99].

Some further insight can also be obtained by considering not just explicit guidelines for what a human therapist should do, but also the underlying processes embedded in art-making which allow people to feel good effects. As noted in Section 1, in art therapy, some positive effects—such as improved self-awareness, self-image, relaxation, and social interactions—are facilitated via processes of self-exploration, self-fulfillment, catharsis, and perceiving belonging [11]. Self-awareness is enhanced by projecting and exploring emotions and experiences, which might be easier to express with symbols than words (described as “refraction”, “dramatic distancing”, or being “once-removed”), and by allowing inference of a person's current state and progression over time; enhanced self-awareness can facilitate healing, via reappraisal. Self-image is enhanced by fulfillment: via opportunities to achieve and to actively take an empowered role in improving one's situation, promotion of cognitive ability (creative thinking), and positive distraction allowing a person to escape from negative rumination to a state of “remembered wellness” and regain an identity not defined by their problems. Relaxation is promoted by catharsis: the release of stress and tensions, engagement in a repetitive physical activity in which the person can freely choose when to start and end, and the subjective nature of aesthetics allowing for many different kinds of “good” result. Social interactions can be improved by feeling a sense of belonging, by being able to share with others, and by being included in a form of expression which is accessible to people of all ages and cultures. Thus, we propose that an art therapy robot can seek to promote self-exploration, self-fulfillment, catharsis, and perceiving belonging.

To promote self-exploration, the robot should sometimes ask a person about their art; in doing so, the person should not be directly interrogated, but rather their emotions can be explored indirectly through the art. The robot can also compile data on a person which can be given if approved to a care giver such as a doctor to make inferences on a person's state and progression. To promote self-fulfillment, the robot could leave the most important parts of the painting for the person to paint, providing opportunities to feel independence, control, purpose, and growth, while possibly scaffolding and adjusting the challenge to the person's skill level. In addition, the robot can refer to the person

as an artist or as creative or skillful, seek to positively distract the person sometimes, and engage the person in creative thinking, asking questions such as “What do you think this looks like?” or “What would you do next time you wanted to paint something like this?” Furthermore, self-acceptance could be facilitated by including positive personalized content; for example, a robot could paint fish for a person with dementia who used to enjoy fishing. To promote catharsis, the robot should not put time pressure on a person, but rather ensure the atmosphere is relaxing; the robot should take over as little as possible the physical act of painting for a person, and should recognize when a person wants to start and end. To promote social interactions, the robot should suggest opportunities, such as showing the painting to others, or displaying a painting somewhere; the robot could furthermore suggest including others in paintings, mention others who have painted similar paintings, or seek to include interesting content in paintings which could lead to conversation.

2.3.2. General Interactions

We have considered the human science literature specific to art therapy to gain some initial ideas. Here, we enrich those ideas by considering general requirements to achieve good interactions with a robot, based on the idea that, although humans can do much automatically without conscious thought, robots need to be explicitly programmed. In other words, positive user experience (UX) is important to realize successful interactions with a robot, but does not automatically result when a system is built; rather, conscious design is required [100]. Here, we discuss properties such as behavior modalities, and how to structure behavior to facilitate general well-being.

A fundamental question is which modalities a robot should use for input and output. The usefulness of speech for art therapy can be expected, as psychotherapy is sometimes referred to as “talk therapy”; verbal and vocal channels allow complex information to be conveyed [101], in a highly salient fashion [102], without requiring a person to look away from art-making and possibly lose concentration [103]. Visual output also is useful, in that numerous streams of information can be shown continuously and simultaneously, additionally to people with difficulty hearing, and a human does not have to wait to hear complete messages from a robot, which could be difficult for users with limited attention. Moreover, tactile interfaces are fundamental, both for operating tools and machines, and for affectionate interactions. Since human art therapists typically utilize multiple modalities, we propose that robot art therapists as well will ultimately benefit from the ability to engage in multimodal interaction.

We also consider the overall problem of facilitating a person’s perceived well-being. In general, a robot can seek to facilitate hedonic and eudaimonic aspects of well-being, i.e., short-term positive feelings, and aspects which should contribute to eventual good feelings over the long run. Short-term feelings can be easier to measure when assessing the usefulness of some intervention, but their relation to long-term happiness is not always clear—especially considering the phenomenon of “hedonic adaptation”, where people after a fortunate or unfortunate event tend to return to the same “set point” or degree of happiness [104]. This mechanism could be beneficial for a person’s ability to survive (e.g., helping people to avoid becoming oblivious to danger, and recover, respectively), emerging from neurochemical desensitization, but reduces the certainty of achieving positive long-term effects when focusing only on short term design properties. Likewise, the impact of focusing only on eudaimonic factors can also be unclear; a person struggling hard to improve themselves might experience much suffering every day, without certainty that their goal will be ever reached. Therefore we believe it is useful to design for both aspects, also in trying to facilitate good feelings toward the artist, the art, and the robot.

In our previous work, we have proposed some designs for facilitating hedonic well-being in an interaction with a robot, based on guidelines that a robot’s behavior should be rewarding, helpful, or inspiring (possibly combining function and playfulness, reactivity and proactivity), clear in regard to its intentions, and carefully executed [105,106]. In addition, some general criteria have been proposed to be important for a person’s eudaimonic well-being [107]—comprising self-acceptance, positive relations

with others, autonomy, environmental mastery, purpose in life, and personal growth—but how to apply them for a robot to support a person’s well-being in a creative application is unclear. We note that these criteria have been embedded in a form of therapy called “well-being therapy” which is intended to help with affective disorders, but this involves a very different scenario from the one we consider, in which a therapist assesses notes on perceived experiences to detect impairments and leads a process of cognitive restructuring [108]. Instead, we argue here that these dimensions of eudaimonic well-being can be related to the qualities positively affected by art-therapy previously noted. Self-exploration allows for personal growth and perceiving purpose in life. Enhancing self-image relates to self-acceptance; as well, being autonomous and controlling one’s environment leads to having a positive self-image. Furthermore, enhancing social interactions entails positive relations with others.

Aside from theoretical prescriptions, the measurement of bodily signals in neuroscience is also contributing new understanding. Korb has summarized some positive effects from feeling gratitude, labeling negative emotions, decision-making, touch, bright light, and exercise [109]. Feeling gratitude resulted in increased levels of dopamine and serotonin, associated with well-being. We note additionally that, not only being thankful for positive experiences, but forgiving of negative experiences, has also been linked with well-being [3]. Labeling negative emotions resulted in higher ventrolateral prefrontal cortex activation and reduced amygdala activity, which relates to reduced perceptions of worry and fear. Making decisions, involving actively selecting and planning actions, especially without striving for perfection, engaged the prefrontal cortex, helped overcome striatum activity related to negative habits, calmed the limbic system, increased dopamine, and led to a stress-reducing feeling of control. Touch resulted in increased levels of oxytocin, serotonin, and dopamine—neurotransmitters associated with well-being—as well as reduced perceptions of pain and social exclusion, which was reported in an fMRI study to affect the brain similarly to physical pain. We also note that warmth is a typical component of human touch, where physical warmth has been linked with perceived psychological warmth [110]. Bright light in the day and exercise also led to numerous positive effects such as boosted serotonin levels.

Based on this, we draw the following conclusions: If a person does not know what to paint, the robot can suggest painting something for which they feel grateful; for negative displays, a robot can ask them to consider forgiveness. A robot should ask a person to describe depicted emotions, although this does not mean a robot should not also make inferences itself based on the person’s art and/or direct signals. As suggested previously, the robot should allow the person to make various decisions. If possible, it would be an advantage for the robot to also be capable of engaging in simplified touch interactions (e.g., recognizing a hug or pat, and reacting appropriately); some previous studies in robotics have proposed such methods [111–114]. Likewise, an advanced system could act to select a painting environment which provides some sunlight and adequate warmth, and if possible should let the person move to get physical exercise (e.g., to fetch paints).

In designing interactions for well-being, another challenge will be how to assess the success of art therapy interactions. For this, a number of instruments have been described, from the Bradburn Affect Balance Scale, Fordyce Happiness Scale, and Satisfaction with Life Scale (SWLS) for well-being, as well as the Friedman Affect Scale and Positive and Negative Affect Schedule-X [3]. For working specifically with elderly with dementia and art therapy, the following instruments have been proposed: Cornell Scale for Depression in Dementia (CSDD), The Multi Observational Scale for the Elderly (MOSES), The Mini-Mental State Exam (MMSE), The Rivermead Behavioural Memory Test (RBMT), Tests of Everyday Attention (TEA), Benton Fluency Task, and Bond-Lader Mood Scale [12]. Thus, many tools exist which can be applied to evaluate robot art therapy interactions.

In summary, we have examined both explicit and implicit guidelines for human art therapists, as well as general guidelines for achieving good interactions, to prescribe interactive guidelines for an art therapist robot.

2.4. Ethical Pitfalls

In addition to considering what a robot therapist should do, we believe it is also important to pay careful attention to what should be avoided. In our previous work, we considered the ethics of a general system that can visualize emotions, e.g., using a computer screen [31], discussing some possibilities for how to conceal a person's emotions or use existing regulation such as the General Data Protection Regulation (GDPR) by regarding a person's emotional state as personal information. From this previous work, we imagine that pitfalls could generally involve physical harm, psychological harm, mistakes, and deception, but this level of detail is insufficient to incorporate into a design. To go further, we need to also consider the specific scenario of painting with a robot in art therapy. To do so, we follow the general pattern of our previous work in first identifying some potential pitfalls, before suggesting some avoidance strategies.

2.4.1. Identifying Pitfalls

Considering that potential problems could involve physical harm, psychological harm, mistakes, and deception, we imagine that physical harm could be caused by the robot, the act of art-making, the person, or others. A robot could hurt a person with its motions, or splash paint or water around, damaging physical property such as the environment or a person's clothes. Art-making could negatively affect a person: for example, making a person with respiratory problems inhale chemicals in paint or varnish, or making a person stand or sit for too long, possibly without drinking liquids. We also consider the possibility that a human could harm the robot or environment: assaults on robots, usually by younger persons, have been reported in the HRI literature [115]; in addition, although incidences are typically rare, assaults on caregivers by persons with dementia, retardation, and (paranoid) schizophrenia have also been described [116]. Lastly, physical altercations or mistreatment—potentially subtle, such as leaving the window to a patient's room open on a cold day—could ensue if peers or caregivers learn as a result of art therapy that a person has negative feelings to them, such as violent obsessions, and is vulnerable.

We also believe that psychological harm could result from scaring a person, damaging relationships, or making them feel bad about their emotions, creativity, or painting skills: Even if a robot is safe, if this knowledge is not adequately conveyed, a person might be afraid of being harmed by or breaking the robot, because people of various cultures can feel anxiety about robots (it is not just limited to the West, where Terminator-like scenarios are commonly considered) [117], people have been killed by robots in the past (<https://www.news.com.au/technology/factory-worker-killed-by-rogue-robot-says-widowed-husband-in-lawsuit/news-story/13242f7372f9c4614bcc2b90162bd749>), and robots can be expensive (e.g., <https://qz.com/1194939/the-us-government-just-gave-someone-a-120-million-robotic-arm-to-use-for-a-year/>). Psychological harm could also result if interpersonal ties are damaged—leading to discomfort or decreased trust (e.g., a family member or friend might wonder if this is really the person they thought they knew)—or from the stigma of having it revealed that the person had engaged in therapy with the robot [118]. In addition, it is unclear if people will feel the same way about a robot therapist as with a human, or if they will feel as if they have been abandoned, to be left alone with a machine [119]. Additionally, Phillips mentioned that patients can initially seek to test art therapists, possibly with disturbing art, and “write off” therapists who are found to be unable to be open, which could result in the therapy failing and perceived isolation [83]. Similarly, therapy could fail because some patients are confused and uncertain, remember failures in previous art classes, and can be reluctant to open up emotionally and wary of revealing vulnerable innermost emotions [1].

Mistakes could involve indiscriminately revealing emotions, recognizing or depicting the wrong emotions, or making mistaken judgments about a person. Inappropriately showing a person's emotions could reveal a mental illness or feelings to others, which the person might not wish to have known. This information could be revealed by accident, or sought intentionally; for example, a care center might seek to check for “problematic” residents. Trouble could arise if the person's

feelings are somehow judged as undesirable, threatening and unjust, and if a person is perceived as weak and an easy target, i.e., crime can result if there is a convergence of a criminogenic strain and circumstances conducive to criminal coping [120]. As noted above, this could result in physical violence or psychological pain. However, we note that such an outcome is not the only one possible; seeing a person's emotions depicted in a painting could also promote understanding and empathy, leading to potentially better treatment [121], or simply not be of interest if emotion visualization becomes common and desensitization takes place [122].

Moreover, if a person feels that the robot could convey their emotions to the wrong person, they might feel concerned, and try to hide their emotions; however, hiding emotions is a difficult task, as emotions can be leaked through facial microexpressions and body language [123], and, because the outcomes could be serious, more worry could be perceived. In addition, although a physical painting might have some advantages over digital media, in that a computer image can be uploaded to the internet where many people could see it, and it might be difficult to remove the image if it exists on various computers, physical objects are harder to store as they require more space, and leaving such a permanent record could act as a reminder of stigma, informing many years later that a problem existed.

Mistaking the emotions that a human is feeling, verbally or in art, could also lead to damaged relationships, from which physical or psychological harm can result, as previously described. Wrong emotions could be indicated intentionally, via the work of a hacker, but we imagine that most cases would likely be unintentional: we expect that mistakes will occur in recognizing emotions, which can be difficult even for people, mostly due to the complexity of the phenomenon. Errors could occur when inferring basic emotions, referents, mixed emotions, or progressions. As an example, emotional signals can be ambiguous. For example, nodding can be positive, indicating agreement, greeting, or thanks; neutral expressing confirmation; or negative conveying irony or emphatic insistence [124]. The robot could infer that a person who has watched a horror movie is afraid of a caregiver. The person could feel happy and afraid at the same time, but be inaccurately described by an algorithm as afraid. The person could also have been sad all day but be assessed as happy at that moment. Errors in algorithms can also act as self-fulfilling prophecies [125]. For example, a person could become angry if an algorithm recognizes them as angry when they are not.

Another possible problem could result not from the robot, but on the observer's side, who might not interpret a painting or its description properly. Such mistakes could result in physical or psychological harm if the misinformation causes a decision to be made. For example, if a person's state is mistakenly inferred to be problematic, unneeded medicine could be prescribed. Conversely, if an extreme state is seen as normal, caregivers could be prevented from providing timely help. For example, negative consequences could ensue if a robot ignores genuine threats, either conveyed verbally or in a person's art; this is complicated because so-called "sublimation", a Freudian term describing tension release through a productive act, can appear negative, similar to threat communications (e.g., as in disturbing art) [83]. Another potential mistake is "contaminating" the personal meaning of an artwork by telling the person what a painting means, as the meaning of a painting for its creator is more important than a therapist's interpretation [1]. In addition, the robot should be careful not to mistakenly exclude a person, by ignoring them or their decisions.

Finally, problems could result from deception, in promoting false relationships in place of real ones, in manipulating emotions, and in creating unrealistic expectations. A person might feel camaraderie toward an art therapy robot which is their partner in painting and receive the impression that the robot understands them; such people might be used to interacting with computers and robots, and it might feel safer to bond with a robot than a human as there is less possibility for rejection and less judgement [126,127]. A danger is that a person might then turn to a robot as a replacement for other humans, even though the relationship will likely not be genuine in the same sense as with humans [128]. Moreover, a robot aware of a person's emotions could be programmed to use its knowledge for some extra purpose, e.g., to influence a person to select some treatment over another one, or to gain money, power, or political support. An added concern is that robots can be perceived to be less accountable

for dishonest behavior than humans [129]. In addition, misleading people with regard to a robot's capabilities or the expected outcomes of the therapy with the robot could result in disappointment and lack of trust, not only in the robot, but potentially also in art therapy itself or any human caregivers involved.

2.4.2. Proposed Solutions to Ethical Pitfalls

In general, to deal with potential problems with physical harm, psychological harm, mistakes, and deception, we propose the detailed solutions below, which are also incorporated as part of our design in Section 3.1. Some proposals, such as regarding safe embodiment and basic recognition capabilities, target robots specifically, as it is understood that human art therapists will behave in a safe and context-aware manner. Conversely, some proposals, comprising patient confidentiality and only acting in a person's best interest, are based on our idea that art therapy robots can also follow codes of conduct for human art therapists (such as the Code of Professional Practice by the Art Therapy Credentials Board (ATCB)). Additionally, both basic and more advanced proposals are presented, as follows.

To prevent physical harm, we propose that a therapist robot should have some capability to avoid and detect problems, and alert and aid when trouble occurs. To avoid hurting a person, the robot itself should be built in a safe manner, and be fully covered, with no exposed sharp, hard, or electrical parts; with compliant joints and quickly cancelable motions; and with a stable base and hardware emergency stops. Moreover the robot's intentions should be easily inferred. This can be tackled by using meaningful motions which are not too fast, and clearly indicating any changes in the robot's motion patterns, potentially verbally. An advanced robot could also detect problems via typical human communication modalities such as sound or vision (e.g., screaming for "help" or frantic hand-waving); moreover, gaze and/or projectors can be used to show where a robot will move next [130]. To avoid damaging property, the environment in which the robot operates should be kept safe and uncluttered; e.g., newspapers can be placed on the floor below a painting to catch paint. A more advanced robot could alert humans if it is being damaged, and seek to avoid damage; in addition, a mobile robot should be able to sense some of its environment (e.g., to avoid falling on a person), and a physics model could be used to predict what kind of damage could occur when planning motions.

To avoid problems coming from the act of art-making, a robot should have some basic ability to disclose potential dangers. Moreover, a contraindication for art therapy could be if it takes the place of another, needed treatment, such as medication; thus, the decision to engage in such interactions should always take place with the knowledge of a person's caregivers. A more advanced design can also have some basic ability to check the suitability of the environment—e.g., temperature, exposure to sunlight or noise, adequacy of ventilation, and presence of furniture if a person becomes tired or faint—and keep track of a person's state, such as how long they have been standing, when they last drank liquids, potentially what medical conditions they have, and how they respond when asked about how they are feeling. Respiratory problems could be taken into account when choosing paints and varnishes. To avoid physical harm from others, a robot should engage in continuous interaction, sometimes asking the person if they feel all right, and alert if there is no positive response. An advanced robot could seek to detect altercations or mistreatment, and even potentially seek to prevent or defuse the situation; in general, we also propose that an advanced robot should also seek to monitor basic signals such as a person's pulse, breathing, temperature, and pose; and to also help in emergencies by alerting medical personnel and potentially even perform basic first aid procedures.

To avoid psychological harm to a person, the robot should clearly explain that it is safe and robust, disclosing its safety features, and should check that a person is willing to paint with it. In addition, the robot should indicate sincere liking of a person through positive behavior, and avoiding negative behavior directed to a person in regard to their emotional states, creativity, painting skills, etc. To avoid perceived stigma, art therapy robots should look like regular robots used for other tasks such as housework or teaching, and should not be associated with illness; furthermore, such

robots can be used by various demographics (children in schools, adults at work, and elderly at care centers) to avoid being associated with any one particular disorder. To avoid being written off, a robot should react positively to a person's art, without showing shock, fear, or disgust [83]. Likewise, to deal with initial negative feelings, the robot should seek to create an atmosphere that is positive and non-judgemental.

To avoid communicating private information to the wrong person, an art therapy robot should practice patient confidentiality. Paintings, physical or in digital form, should be stored safely, in locked safes or computers not connected to the Internet. Processing can be on-board, potentially communicating only basic information if needed, via encryption, and conducting authentication. Cloud services might not be optimal if there is a risk of data being lost, corrupted, or stolen. Sharding (breaking apart files and spreading them over multiple computers) can be used instead of working with whole files or data blocks. For encryption, popular current methods incorporating public key cryptography such as RSA (Rivest–Shamir–Adleman) and PGP (Pretty Good Privacy) are expected to be breakable by quantum computing able to run Shor's algorithm and find the two primes of a semiprime in quadratic time; some post-quantum alternatives include lattice and code-based methods such as NTRU and McEliece (<https://www.zdnet.com/article/ibm-warns-of-instant-breaking-of-encryption-by-quantum-computers-move-your-data-today/> and <https://www.technologyreview.com/s/420287/1978-cryptosystem-resists-quantum-attack/>). In addition, information that is made available to a robot can also be limited, especially if the robot is part of a team of caregivers; for example, the robot probably should not need to know which medications a person is taking. Moreover, a robot could generate dummy paintings implying that a person has felt only positive emotions. For higher security, private rooms can be used, without cameras and with closed windows and curtains (e.g., preventing drones from spying); art-making could also be conducted in virtual or augmented reality.

To avoid problems recognizing emotions, the robot should never act in isolation, but always ask for feedback and confirmation from the human. Additionally, more advanced mechanisms for recognizing and expressing emotional properties (e.g., mixed emotions, referents, timing, and polysemy) in visual art will facilitate accurate communication; for example, responsible use of emotions by a robot might include ensuring that displayed emotions are clearly tied to a referent. To avoid mistaken judgements by a robot, we also suggest, in line with our idea that art therapy robots should not be used to judge people, but rather act as partners, that robot behavior should also be logged, along with reasons for any behavior, for transparency. In addition, the robot could alert a human care giver when a person behaves in an extreme manner. For this, it could be possible to only act if it is known from the outset that there is a risk, or if the robot is highly certain of an inference, as in significant hypothesis testing, where a null hypothesis is rejected only if it would be highly unlikely given observed data (e.g., with a probability less than one in twenty, two hundred, or one thousand). It could also be possible to log a person's verbal description of their artwork for a human therapist to review later. To avoid contaminating the meaning of a painting, the robot should ask about its meaning, rather than telling the person what it means.

To avoid potential deception, human therapists should be prioritized when available, and robots only used either to help humans to work more efficiently or when no human is available. It should be made clear that current robots are not capable of forming relationships in the same way as humans. Advanced robots can also seek to promote social ties. With regard to the potential for manipulation, robots should again practice confidentiality, and there should be no ulterior financial incentive to manipulate a person's emotions from the perspective of the caregivers, or partners offering the robot; the only goal of using the robot should be to better help more people to achieve a better state of well-being, by improving quality and quantity of care, and reducing the workload on human therapists. Furthermore, robots' capabilities and expectations for therapy sessions should be made clear to avoid a person feeling disappointed.

By following such prescriptions for what a robot should not do, we believe better interactions will result.

2.5. Emotions

Emotions, as noted previously, form a vital portion of the fabric of art therapy and art-making; therefore, we must examine the literature to determine how a robot can engage appropriately on an emotional level. We note that, by emotion, we do not mean that our focus is on whether artwork is aesthetically pleasing; likewise, we are not concerned with the question if robot emotions will ever be exactly the same as human emotions. Our goal is not to impress people with the beauty of a robot's art or create an identical replication of a human, but to allow robots to engage in emotional interactions with people to support their well-being. Below, we consider modeling, expressing (abstractly or symbolically), and recognizing emotions.

Modeling of emotions has been complicated by difficulty in defining the term and a vast number of proposed models. In engineering, various work is being conducted on emotions using various labels such as affective computing or semantic analysis, with discrete, continuous, or hybrid models, each with advantages and disadvantages. The discrete model does not encapsulate intuitive interrelationships between emotions, such as that "contentment" is closer to "happiness" than "anger", and various arbitrary numbers of categories are used. Likewise, dimensional models can be difficult to imagine (it can be easier for lay persons to imagine "angry" than "low valence, high arousal") and can have trouble capturing some emotions (e.g., "surprise" via dimensions such as valence and arousal). Some hybrid models have also been proposed such as Plutchik's wheel [131], which seeks to place similar categories next to one another in a continuous graph, but this approach suffers from the problems of discrete models; without dimensions such as valence and arousal, the meaning of distance becomes unclear (e.g., what does it mean to place "disgust" opposite to "trust"?). Thus, current approaches also do not take into account the complex properties we described in Section 1 in regard to mixed emotions, referents, timing, and polysemy. In the current work, we suggest the usefulness of using both discrete and continuous models simultaneously: discrete models for working with participants in experiments, continuous models for robot programs, and a map between them. Moreover, we believe in the value of considering advanced properties of emotions for interactions in which emotions are crucial, such as therapy.

We turn to how emotions can be expressed. Various hints can be found in the literature for how to convey emotions through basic abstract properties of a painting, such as color, lines, and composition. From a dimensional perspective, valence can be conveyed by color hue [132] and brightness [133]; shape/line curvature [134] and symmetry [135,136]; and harmoniousness of the composition (key points aligned along lines dividing the composition into thirds, as in the rule of thirds [137], and alignment with the canvas [138]). Arousal can be conveyed by color (warm vs. cold) [132], and combinations (complementary or discordant, vs. analogous) [139]; as well as line orientation (horizontal vs. diagonal) [140]. We incorporate these ideas into our design and prototype described in the next section.

In addition to abstract painting, recognizable content can be used to convey emotions. A highly useful study by Machajdik et al., in addition to describing machine learning features which can be used to classify emotions in images, proposed detecting the presence or absence of people via faces and skin color, and described a need for recognizing other semantic information [141]. We agree that, despite many possible confounders, there will be some typical shared emotional meanings which are perceived in symbols. For example, a heart or gravestone could be painted to express positive or negative emotions, respectively. Some work has been done on automatically recognizing such symbols; for example, local self-similarity descriptors have been used to detect hearts [142], based on the idea that humans can detect symbols expressed in various ways even if there is no simple similarity in typical photometrics like color, edges, or intensity. Furthermore, classifiers trained on photos have been used to detect objects in paintings [143]. In addition, AutoDraw, based on data acquired through an

online quiz called “Quick, Draw!” [144], recognizes symbols users sketch, from a set of categories [145]. What is unclear from the perspective of expressing emotions is which symbols will be able to convey which emotions, in artwork and paintings, as is addressed further in Appendix D.

In addition to expressing emotions, a robot can also seek to recognize a person’s emotions. Some readers might wonder if an art therapy robot truly needs to be able to recognize emotions, or if the robot could not merely imitate the colors and lines that the person is using, or rely entirely on asking the person what they are feeling. We believe that emotion recognition is useful: by recognizing the emotions behind a painting we expect a robot can provide a more creative and more interesting performance, introducing new colors, lines, and content—especially when a person’s painting skill is low. Furthermore, a person’s response might not be accurate, as sometimes people do not even know how they are feeling. Additionally, in the case where a person cannot or does not wish to make art themselves, the robot might have to rely on signals other than in the painting to infer their emotions. Various sensors, from cameras and microphones to electrodermal activity sensors, and brain machine interfaces, can be used to infer emotions, from signals such as facial expressions and speech [146–148].

In particular, we believe that in art therapy, analysis of a person’s spoken words, also about the artwork, will be highly important. Conversational devices such as Google’s Assistant, Apple’s SIRI, and Amazon’s Echo are already engaging in conversations with people in homes; Google Duplex, based on a recurrent neural network (RNN) using TensorFlow, has demonstrated some excellent performance in conversing with humans, e.g., in calling a restaurant to make a reservation [149]. To extract important information about emotions from verbal content, some challenges include anaphora resolution, word sense disambiguation, aspect extraction, and named entity recognition (pertaining to referents); subjectivity analysis (which would seem to be useful for therapy, in recognizing which utterances are revealing, perhaps for summarizing sessions); as well as detection of insincere text incorporating sarcasm and humor [150]. There are also many examples of robots designed to engage in speech-based interaction: for example, Kismet, capable of detecting affective vocal displays tailored to the robot’s child-like appearance [151], and Nico, which can learn new words by generating definition trees comparing parsed utterances with sensor-detected entities [152].

A demerit to typical modalities such as face expressions and speech is that they are relatively easy to control, e.g., faking a smile or a cheerful speech response, and might not work with some groups, e.g., with autism, strokes, or paralysis. Various sensors could be used to avoid such problems. For example, Rani et al. proposed a robot which can detect a person’s stress via wearable sensors [153]. Here, we suggest that BMIs and thermal cameras could also be useful, which we feel possess some advantages such as recognition power in the case of BMIs and remote sensing for thermal cameras, but have not received as much attention yet as other sensors. BMIs have been used to detect positive or negative valence via higher activations in the left or right frontal lobes, respectively [154]. Facial temperature has previously been used to infer arousal [155], and also to detect Ekman’s six basic emotions [156].

Thus, currently, there exist various models, hints for expression, and techniques for recognition, with a few noteworthy gaps, such as in regard to the emotional meaning of symbols, as well as methods for working with mixed emotions, referents, progressions, and polysemy.

2.6. Creativity

In addition to emotions, art-making, also within the context of art therapy, is an inherently creative process. A robot should not paint randomly, or always in the same fashion, disregarding the human’s emotions, which could feel boring, unhuman and hard to relate to. Rather, an art therapy robot should convey appropriate emotions, in a creative manner which stimulates and engages.

Here, we describe some recent work on artificial creativity, where great strides are being made. Computers are composing songs [157], authoring short stories [158], generating news articles (e.g., <https://automatedinsights.com/wordsmith> or <https://narrativescience.com/Platform>), writing

movie scripts [159], and creating computer games [160]. Some artworks could be described as passing a Turing test, in the sense that they are sometimes believed to be created by humans or even rated higher than actual human-made creations: for example, poems [161] and particularly relevant to the current proposal, visual art on a computer screen [162].

A central question is how a robot should engage in a creative process. Some inspiration can be found in computational theories that are beginning to elucidate how creativity is exercised in humans. For example, some recent work based on the concept of the “adjacent possible” [163] has succeeded in predicting some laws of creativity such as Heap’s Law and Zipf’s Law, that the rate at which new artifacts are created is sublinear, and that rank and frequency are inversely related within creative spaces, within a simplified context (the Polya’s urn problem) [164]. Basically, such work involves having some model of what is possible and some way to reach adjacent terrain. We believe this basic formulation is reflected in general in current approaches for artificial creativity, which tend to comprise two components, some kind of prior model or data, along with some mechanism of introducing noise.

One powerful “black box” mechanism for creation involves using a generator-discriminator combination. For example, in the work mentioned above by Elgammal and colleagues, a “Creative Adversarial Network” was built, in which feedback from the discriminator was modified to rate art both in terms of goodness and style; confusing the discriminator with style is used to assess creativity. The goal was to generate art which, in accordance to guidelines by Berlyne and Martindale, would create a suitably high arousal potential, being somewhat creative, but still adhere to typical notions of what is aesthetic. A drawback to such approaches is difficulty in explaining why decisions were made. As in the third level of DARPA’s categorization of artificial intelligence (where the three levels refer to hand-crafted rules, statistical inference, and explanatory AI), there is an increasing trend toward transparency and ability to explain decisions, which is especially important if a robot is supposed to engage in a relationship of trust with a human [165]. For example, in art therapy, if a robot generates responsive art which could be interpreted badly, it should be able to clarify its intentions, to ensure that a human’s feelings aren’t hurt. One possible approach to realize such a transparent creative system could apply other statistical approaches which generate human-readable rules, such as trees or boosting, although obtaining stable classifiers which do not vary with small changes in training data can be challenging in practice. Another possibility could be to use the discriminator not just on generated artworks, but to identify models which are more creative. For example, creativity has been assessed in one work in terms of the number of different solutions which can be generated to solve a problem [166], thus perhaps progressively more creative models could be trained.

Another fundamental approach is to use some outside data to “seed” artwork with creativity. For example, Colton et al. described an AI which generated poetry from online news articles; its creative decisions were explained in a report it automatically created about each poem [28]. We believe such an approach to be highly useful. A demerit from the perspective of art therapy could be that a robot’s artwork will likely not be generated immediately, but rather people will watch the artwork being generated. Thus, it could be beneficial if people could understand what a robot was doing without waiting until the robot has finished. Moreover, people might not bother to read descriptions about a robot’s painting. Alternatively, if read out loud, people might not wish to wait through a long report. We propose that it would be useful for a system to be able to explain creative decisions in a more natural fashion, such as verbally in a conversation. In doing so, the robot could proactively issue statements and questions to seek to assess a person’s interest and knowledge levels (knowledge tracing), arouse curiosity, and summarize or expand on topics as desired, adjusting the detail of its words to take into account a person’s knowledge and interest level; in this way the robot could also immediately halt its explanations when appropriate.

We believe that some possible topics of interest here include how to find a good balance between questions about a person’s artwork and comments about its own, how a robot can detect a person’s engagement or boredom as feedback, how it can quickly and accurately infer interest and knowledge levels, possibly from a cold start where no data are available (i.e., zero acquaintance), and, although we

concern ourselves primarily with a dyadic scenario in the current article, how to effectively explain to groups when interest and knowledge levels of individuals differ. Additionally, we have assumed that it is important for the robot to explain why it paints the way it does but other questions are also possible. Fox et al. listed the following questions which explainable AI systems should be able to answer: Why did the robot not do something else? Why was a robot's behavior good in terms of some particular criteria? Why can't the robot carry out some particular task? Why is some replanning/reconsideration required or not required [167]?

In summary, a vast literature exists, spanning robots in therapy and art, interactive strategies, and mechanisms for emotion and creativity, from which a large number of prescriptions can be made.

3. Design for an Art Therapy Robot

3.1. Requirements and Capabilities

In the previous section, various prescriptions were made, which required merging and structuring to form a design. Thus, we revisit requirements, goals and problems, as well as proposed solutions, identified in the previous section, grouping related concepts. For example, relaxation and catharsis were included in the category for hedonic well-being, as they refer to temporary good feelings. This yielded seven requirement categories and six capability categories. Requirements categories, explained below, are as follows: R1 Co-explore, R2 Enhance self-image, R3 Improve social interactions, R4 Please, R5 Engage Emotionally, R6 Engage Creatively, and R7 Avoid Pitfalls.

- R1 Co-explore. The robot should investigate the meaning of the person's art with the person, showing attention: expressing inferences about the person or their artwork in painting and verbally, and asking questions to confirm and to encourage the person to reflect. The robot can also track the state of the person's artworks over time.
- R2 Enhance self-image. The robot should be positive, accepting and encouraging the communication of emotions. If sharing a substrate, the robot can also leave the most important parts of the painting for the person to paint, providing opportunities to feel independence, control, purpose, and growth, while possibly scaffolding and adjusting the challenge to the person's skill level. The robot can seek to include positive personalized content tailored to the person in its painting.
- R3 Improve social interactions. The robot can suggest including others in paintings, mention others who have painted similar paintings, or seek to include interesting content in paintings which could lead to conversation.
- R4 Please. To promote a general perception of hedonic well-being, the robot should help the person to feel good about engaging in art therapy with the robot, by behaving in an enjoyable and likeable way. For easy good communication, the robot can offer a familiar interface such as humanoid behavior and capable of multimodal interaction. To be liked by the person, the robot can show empathy and match a person's emotions, although it should not show negative emotions toward a person or their art; and be positive, showing sincere liking, as praise can be given for a person's creativity and skill, even for negative depictions. Emotion expression can be made to be large and meaningful, and express sincerity by being clear, e.g., ensuring that referents are clearly conveyed. Creativity can show interesting variation within a stable core. Suggestions, delivered proactively, can invite interaction and clearly convey interactive affordances and a robot's intentions. The robot can also infer when the person wants to end the interaction.
- R5 Engage Emotionally. The robot should seek to infer emotions embedded in a person's artwork, and can also seek to infer a person's emotions directly. Basic emotions can be conveyed abstractly based on heuristics, or via symbols such as a person's face. The robot can also seek to infer and convey complex phenomena such as mixed emotions, referents, and progressions.
- R6 Engage Creatively. The robot should be able to make basic creative choices and discuss these with a person through conversation.

- R7 Avoid Pitfalls. The robot should avoid pitfalls including physical harm, psychological harm, mistakes, and deception.

In addition, six categories of capabilities are defined as follows:

- C1 Infer. C1.1 Emotions from painting or C1.2 emotions directly from a person. C1.3 Speech, when person wants to end the interaction, asks a question, or answers. C1.4 Problems.
- C2 Paint. Express C2.1 Matching emotions. C2.2 Positive emotions. C2.3 Creativity. C2.4 Complex Emotions.
- C3 Talk. C3.1 Ask. About person's painting and permissions. C3.2 Greeting C3.3 Explain. C3.4 Suggest. Social interactions. C3.5 Be Positive. C3.6 Alert.
- C4 Track State. (If permitted) Store a record for a care giver.
- C5 Offer Familiar Interface.
- C6 Be Safe.

The relationship between our proposed requirements and related solutions is summarized in Table 2 .

Table 2. Design for art therapy robot: relating requirements to capabilities.

Requirement	Solution
Art therapy. Co-explore. Enhance self-image. Improve social interactions. Please.	Infer, Paint, Ask, Track State. Be Positive. Suggest. Offer familiar interface, Paint (Match, Creativity), Suggest, Be Positive, Infer, Greeting
Engage Emotionally.	Infer/Paint
Engage Creatively.	Paint creatively, Explain.
Avoid Pitfalls. Avoid Physical harm (from robot) (from art-making) (from others) Avoid Psychological harm (intimidating) (making people feel bad) Avoid Mistakes (Indiscriminately revealing emotions) (Recognizing/showing the wrong emotions) (Mistaken judgements)	Infer problems, Alert Be Safe (safe design, self-diagnostics, transparency about activities and intentions; physics model, capability to detect paints, water, substrates, etc.) Explain (disclosure of potential dangers), Infer (checking for risks) Infer (detect negative human interactions), Alert (try to prevent/defuse) Explain (disclose safety) Praise (positivity toward the person)
Avoid Deception (Taking the place of humans) (Manipulating emotions) (Disappointing)	Be Safe (patient confidentiality, safe storage of paintings, secure system) Paint Complex Emotions (use rich expressive model (e.g., mixed emotions, referents, timing)), Ask (give feedback and confirm), Be Safe (secure system) Explain (transparency, collaborative), Track State (not taking the place of humans) Explain (prioritize/assist human therapists, transparency), Suggest (promote social ties) Be Safe (confidentiality, no financial incentive) Explain (ensure expectations are clear)

3.2. Simplified Implementation Example

It could be challenging to implement the design with only general guidelines. To ensure the design can be implemented, and help others to do so, we describe an example of a simplified implementation. For this, a medium fidelity prototyping strategy was adopted, which balances insight into how a system will be perceived and flexibility and ease of development [168]. Below, we describe the interface, inference, painting, and speech.

3.2.1. Safe, Familiar Interface

A robot was required; various robots were available at our lab, including Turtlebot, Thymio, ARDrone, Nao, and Baxter on Ridgeback (hereafter Baxter). For art therapy, we desired a platform

which would be highly safe and have a familiar interface—with a humanoid appearance that would make it easy to recognize as a therapist—and rich multimodal interactive capabilities, also facilitating development. Therefore, we chose the Baxter robot shown in Figure 3, which is a safe platform intended to operate near humans; all actuators are equipped with springs, and operate at low speeds (the base’s maximum speed is 1.1 m/s). Moreover, Baxter is a humanoid, with a display showing a face capable of showing various expressions (for which OpenCV was used), speakers to issue speech utterances, long seven degree-of-freedom arms able to reach various points on various sizes of canvases without getting in a human’s way (a reach of 1.2 m), and an omnidirectional base. To sense, the robot also has a camera and 13 sonar sensors around its head, infrared range sensors and cameras in its hands, force sensors and accelerometers in its arms, a laser and IMU in its base, and a microphone. In addition, as the robot is adult size, it could be possible for people to easily imitate, and possibly feel a connection to, the robot. We note that one disadvantage of using Baxter in studies with some user groups, such as dementia patients, is that the robot could be perceived as threatening, as it is large and heavy (approximately 100 cm × 80 cm × 180 cm, with a weight of approximately 210 kg), and its bent arms with the elbows upwards could be perceived as resembling an insect such as a praying mantis or an arachnid; in this case another robot could possibly be used.

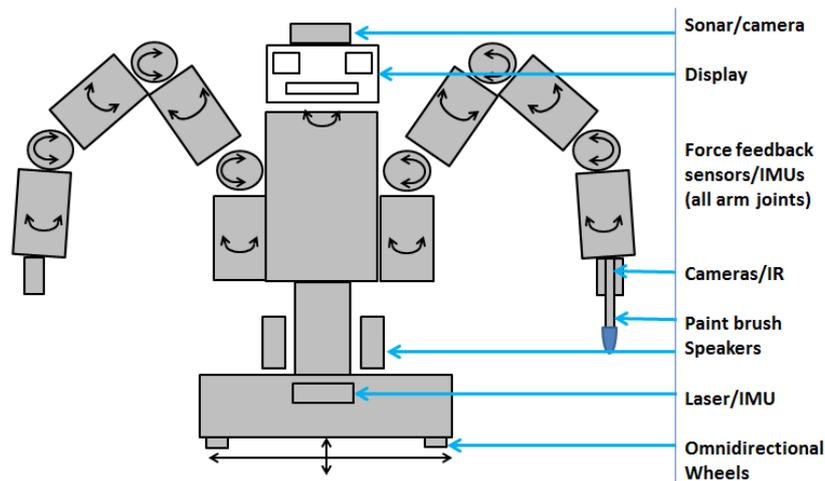


Figure 3. Baxter, showing sensors and actuators.

3.2.2. Inference

Inference addressed detecting emotions with a BMI and recognizing keywords spoken by the person. Russell’s dimensional model was used to model a person’s emotions. To recognize a person’s emotion, we used a typical 14 channel wireless EEG with a sampling rate of 128 SPS; the Emotiv Epoc+. Electroencephalography (EEG) signals were obtained from four channels (AF3, F3, F4 and AF4) on the frontal lobe, and filtered into three bands using the Remez exchange algorithm, which iteratively finds a filter minimizing the maximum error in the desired frequency ranges [169]: Theta (4–8 Hz), Alpha (8–12 Hz) and Beta (12–30 Hz). Mean log-transformed brain wave power values were computed for each of the four channels and three bands by extracting the spectral densities via Welch’s method [170]: this involved averaging periodograms obtained using the Discrete Fourier Transform in Equation (1) on overlapping segments of the signal to reduce variance. These values were used as features by Support Vector Machines (SVMs) in a one-vs.-one arrangement with a Radial Basic Function (RBF) kernel to infer degrees of arousal and valence.

RBF SVMs can be used to classify using Equation (2), where sgn is the sign function returning 1 or -1 , $y_i \in \{1, -1\}$ are supervised labels, a_i are learned weights, x_i are training vectors, x_j is a test vector (e.g., data signals representing a person’s current emotion), $exp(x)$ is the exponential function calculating the value of e to the power of x (the portion of the Equation (2) within the exponential

function is the kernel), b is a bias term, and γ is a parameter greater than zero which represents the influence of support vectors: a high γ will result in low influence of each support vector, tending to result in a classifier with high bias and low variance. Classification values were then sent wirelessly to a desktop program which used a simplified model to associate emotions with visual design elements.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \quad (1)$$

$$f(\mathbf{x}_j) = \text{sgn}[\sum y_i a_i \exp(\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) + b] \quad (2)$$

To hear keywords spoken by a person, CMU Pocketsphinx (<https://cmusphinx.github.io/>) was used for English (we have also experimented with a Swedish version using Voxcommando (<https://voxcommando.com/>)). The robot recognizes keywords indicating that the person is done painting, conveying the emotions that a person's painting shows, and asking the robot about its painting.

3.2.3. Painting

We focused on the basic simplified case in which the robot and human paint on separate substrates. To find heuristics capable of expressing basic emotions, we obtained insight from two professional artists. They helped us by pointing out useful sources in the literature, drawing sketches as examples of how to convey emotions through painting as in Figure 4, suggesting we consider the medium (we had been using canvas, but watercolor paper facilitated aesthetically interesting blending and was found to be useful especially for conveying relaxed emotions), and highlighting the importance of introducing variations into each painting the robot should produce. Together with them, we derived some initial heuristics, shown in Figure 5. A highly simplified strategy was adopted to transform recognized emotions into a plan for a painting. Recognized emotions were transmitted as a pair of real numbers. The real numbers were binned into one of six categories for arousal, and the same for valence, where each bin corresponded to a model for a painting.

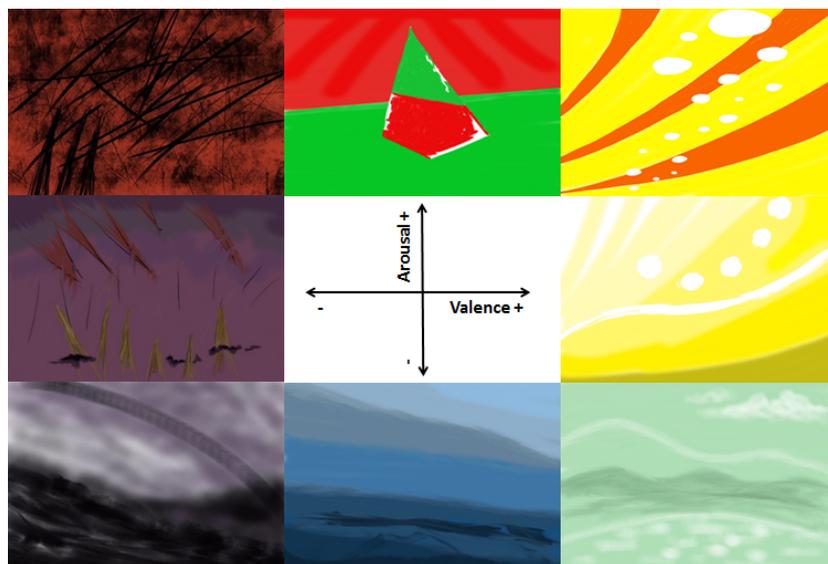


Figure 4. iPad sketches by an artist, illustrating some possibilities for expressing emotions visually. All images ©2017 Dan Koon. Used here with permission.

To add some simplified creative aspect and avoid producing the same painting multiple times, some randomness was added, which affected angles; number, size, and position of shapes; and colors, within the allowed range of each model. Based on the computed composition, commands were sent to

the robot’s internal computer to direct it to paint using a paint brush on its left arm and sponge on its right arm. Along the way, we also noted some unintended effects based on comments from colleagues: One blue background the robot generated felt aroused, an orange/yellow background felt relaxed, and black was felt to express positive valence. We believe this was caused by high contrast when the background was not completely filled in, letting the white of the paper come through; long horizontal strokes used to fill in the backgrounds and low contrast; and a personal color preference, respectively. Not filling in the backgrounds perfectly, while generating some confusion, created some complexity which could be aesthetically pleasing. Using the horizontal strokes for each background also was a shortcut used for practicality. Thus, we recommend designers be aware for the potential conflicts between clarity of communicating emotions, aesthetic complexity, and shortcuts used for practicality. Some examples of abstract paintings expressing basic emotions are shown in Figure 6.

Valence			Arousal		
Feature	High	Low	Feature	High	Low
Color (Brightness)			Color (Warmth)		
Line/Shape (Curvature)			Color (Comp./Analogous)		
Line/Shape (Symmetry)			Line direction		
Composition (Rule of 3rds)					
Composition (Alignment to Canvas)					

Figure 5. Heuristics for conveying basic emotions through abstract painting.

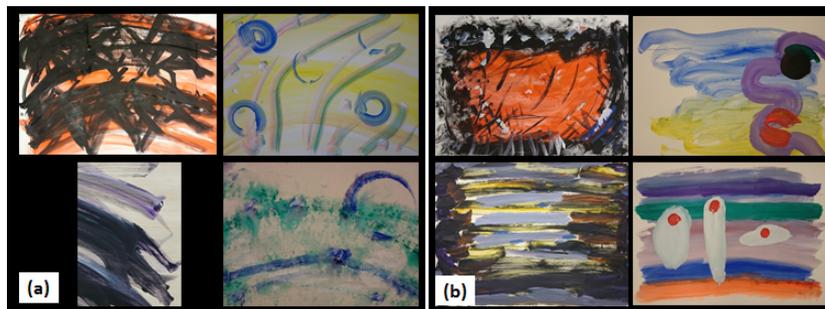


Figure 6. Examples of (a) abstract paintings by our robot, compared with (b) paintings by an artist. In each, the top left painting represents the angry quadrant, the top right represents the joyful quadrant, the lower right represents the relaxed quadrant, and the lower left represents the sad quadrant.

We also investigated how more complex emotions could be conveyed. The easiest way to convey such information could be through a written description, such as meta-information associated with an image, but a person might not read such a description. Referents could be conveyed by painting a human who shows the referent somehow via gaze, pointing or holding some object; or by a thought

bubble or arrow. Mixed emotions could be conveyed by painting a human and using different body parts such as the eyes, mouth and hands as channels (e.g., like the female subject in the Mona Lisa, with her slightly sad eyes and slightly smiling mouth); another mechanism could be to somehow connect emotion displays (e.g., by placing them side by side, such as the Sock and Buskin masks of comedy and tragedy, the Janus head looking into the past and future, or the yin-yang symbol expressing a balance between positive and negative aspects). Progressions could be shown by using the tendency of people to read from left to right, or right to left, in a culture (e.g., comic book style). A polysemy model could be used to select clear symbols. Some examples of symbolic paintings expressing complex emotions are shown in Figure 7. (Additionally we also entered some paintings generated by this prototype into the international robot art competition. The paintings were judged partly by a panel of judges and partly by the public, and received a sixth place award; a comment from the judges was “If this body of work was exhibited at a gallery and I was told that the artist aimed to capture emotion through colour, composition, and textures—I would buy it (says one of our professional judges). The bold brush strokes, cool or warm templates to match the emotional quality expressed, it all made sense—but felt alive. Loved them”).

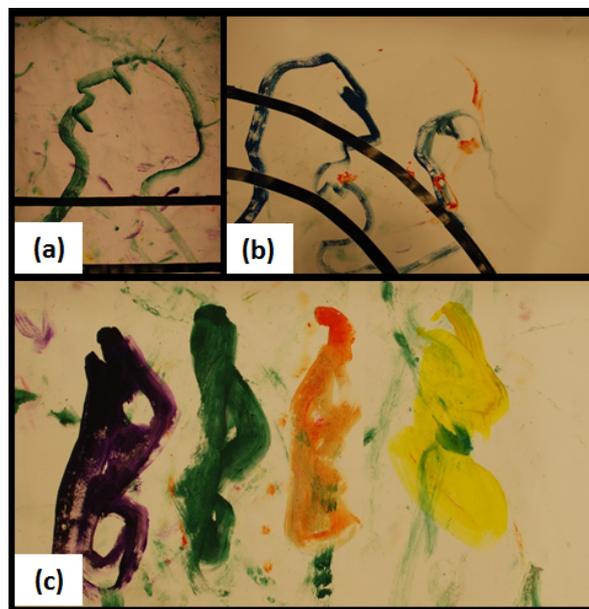


Figure 7. Examples of symbolic paintings by our robot. (a) Expressing basic relaxation (b) Mixed emotions: seeking to express anger (face on left), fear (face on right); sadness (via blue color); and overall negative valence (descending mood lines) (c) A progression from miserable on the left to joyful on the right, using highly abstracted faces.

3.2.4. Speech

For speech playback, *soundplay.node* (Festival) was used. The robot starts the interaction with a greeting, followed by an explanation and suggestions, then loops while waiting for a person to say they have finished painting. A timer is used for the robot to occasionally ask the person about what they are painting or make a comment about its own painting. During the time when it is not speaking, the robot also responds to questions about what it has painted, explaining creative decisions, and if the robot hears a request for “help”, it simulates sending an emergency call to a care giver. At the end, the robot praises the human and says goodbye. If the robot received permission, it takes a photo of the scene with the finished paintings.

Below, we present some examples of the robot’s utterances:

- “Hello, welcome to art therapy.”
- “My name is Baxter and I am learning how to interact with humans. Please don’t be disappointed if I don’t understand what you say.”
- “Do you feel like doing some painting today? It’s fine to say ‘no’.”
- “Is it okay if I let . . . know that you have painted with me today, and take a photo of your painting to show them?”
- “Also, do you know that you can stop my arm at any time, or push the emergency button to make me turn off? I’m completely safe.”
- “Let’s get started! Please paint whatever you like, and I will do the same”
- “What are you feeling right now?”
- “Are you feeling . . . ?”
- “Thank you for asking about my painting. I decided to . . .”
- “You’re done? Okay, then I’m done, too.”
- “Thank you for painting with me, and showing me your wonderful artwork, goodbye!”

Although highly simplified, we feel the prototype illustrates some of the promise of robot art therapy.

4. Discussion

In conclusion, we have explored the theoretical design space for an art therapy robot, suggesting the importance of taking into account the rich literature in human science and engineering on various relevant topics from robots used for therapy and art, to potential strategies for interacting, and mechanisms for expressing emotions and creativity.

- Art therapy robots. We have motivated why art therapy robots would be useful, and described an apparent gap in the literature in regard to drama, writing, and gardening therapy for robots.
- Therapeutic interactions. We have suggested the usefulness of a humanistic, “responsive art” approach as a starting point for an interaction strategy for art therapy robots, comprising concepts such as “matching” and “distracting”.
- Emotions. We have compiled with the help of the artists a list of heuristics for autonomously generating abstract emotional art based on simplified properties of color, lines, and composition. We have reported on some symbols which appear to strongly convey emotions, proposing the importance of one symbol in particular, a painted human face with various expressions, as a familiar and powerful symbol. Furthermore, we have highlighted a perceived gap between our understanding of emotions in human science and what is currently typically being addressed in engineering studies, in terms of mixed emotions, referents, timing, and polysemy, and suggested how such emotional characteristics can be considered or conveyed in art therapy.
- Creativity. We have discussed some issues in artificial creativity, and proposed that an art therapy robot should be able to discuss creative choices with a person through conversation.
- Ethics. We have identified some potential ethical pitfalls for an art therapy robot and proposed solutions for avoiding them.

Based on our arguments, we have proposed a design for an art therapy robot, also discussing an example of a simplified prototype implementation.

We believe that this will help to guide some next steps in this area, although we note several important limitations. Robot art therapy relates to a huge number of research areas, which cannot be comprehensively surveyed in a single article; rather our intention was to discuss some central studies and present an exploratory overview, to gain insight into the theoretical foundations for robot art therapy. As such, there are risks of bias, and areas in which the body of evidence is sparse. For example, robots typically use large datasets to learn, but descriptions of how human art therapists respond

with visual feedback are currently few. Furthermore, our suggestions are limited by lack of certainty regarding concepts such as emotions, creativity, and well-being, which are still being studied in human science. Moreover, we do not suggest that our design will allow robots to replace human art therapists, which we believe is not desirable and would also be currently impossible from the perspectives of conversational ability and understanding of humans; rather the goal in the current article is simply to act as a foundation toward setting up some first simplified interactions expected to be beneficial for humans, while ensuring that expectations are realistic, and permitting some suspension of disbelief, although we expect problems regarding overly high expectations of a robot's abilities will fade over time due to accustomization. In addition, we believe that, even if a human therapist were required to analyze data from sessions, make inferences about progression or regression of the patient, and plan how to structure next sessions, partner robots could still be useful, e.g., in potentially providing some care to people who cannot currently receive care due to lack of resources, facilitating self-exploration in an engaging, non-judgemental way as a partner, being available at possibly any time as robots do not need to sleep, and saving some time for human therapists, such as time spent traveling to, attending, and documenting sessions.

Future work will advance both our theoretical and practical knowledge. Below, we give a few examples.

- Art therapy robots. Theoretical work will extend our design to work with other forms of therapy such as music therapy.
- Therapeutic interactions. Interaction strategies will be refined, e.g., by clarifying how art should be used to provide therapeutic feedback, possibly through preparing datasets of patient art and responsive art from human therapists.
- Emotions. We will get a better understanding of how to convey complex emotions, also using symbols, and we will tackle some questions which emerged in our work such as: Will abstracted symbols be perceived more as conveying emotions, whereas realistic symbols will be perceived more as semantic or referential?
- Creativity. Various questions can be tackled, such as: How could an art therapy robot personalize creative artwork generated for a person?
- Ethics. Concerns for other forms of art or therapy should be addressed, as well as legal questions. For example, if an art therapy robot errs, who is at fault: the robot, the makers, or the entity offering its services? Moreover, who will own the copyright for the robot's generated artwork if it is generated in a therapy session for a human?

Practical work in general will involve conducting interactions with various users, and assessing various design capabilities and assumptions. By exploring how robots can contribute to our well-being in emotional and creative interactions, our intention is that the resulting knowledge will also positively affect acceptance of robots in human environments, provide some insight into which facets of human activity are unique or also performable by robots, and—in the sense that robots can provide insight into human nature [171]—possibly even move a small step toward gaining a better understanding of the complex human phenomena of emotions, creativity, and well-being.

Author Contributions: M.D.C. conceived of and wrote the article. M.L.R.M. contributed some knowledge on emotion recognition with a brain-machine interface, and together with M.C., supervised a student project on robot art, which is briefly described in the current article.

Acknowledgments: We thank various people who helped us, including the artists, Dan Koon and Peter Wahlbeck, as well as Fundación Ageing Social Lab and Macrosad, who helped us to have some conversations with prospective elderly users, with and without dementia. This research was funded by the Swedish Knowledge Foundation (Sidus AIR No. 20140220), also for publishing in open access, and some travel funding was received from the EU REMIND project (H2020-MSCA-RISE No 734355).

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A. Emotions

Emotions are key to art and art therapy [6]—so much so that the purpose of art has been defined as being “that feelings . . . can be transmitted with all their force and meaning to other persons or to other cultures” [5]. “Emotion” in humans refers to a complex psycho-physical phenomenon, involving subjective experiences referred to as feelings, somatic symptoms, expression via affect displays such as smiling, and cognitive appraisals [7]. Affective space can be partitioned for simplicity into some discrete categories such as “happy” or “sad”, capturing the way people typically describe emotions [172,173], or along a few continuous dimensions such as valence (positivity) and arousal, characterizing interrelationships between emotions [174,175]. Here, the term emotion is used interchangeably with “affect”, but we differentiate emotions from “moods”, attitudes and “sentiments” which are also considered to be more long-term [176]—although what exactly constitutes short- or long-term is not clear, and we believe that these phenomena are highly intertwined; furthermore, the latter two concepts focus on the dimension of valence rather than on other dimensions such as arousal or dominance. In the context of art-making, we note that an artist’s emotions are not necessarily the same as the emotions which are intended to be perceived in the art, or the emotions that an observer actually feels. For example, a person might have fun painting an angry painting, which might be considered boring by an observer who has seen many similar paintings. Nonetheless, we consider that in art therapy a goal is typically to explore one’s feelings, so there will likely be some correlation between a person’s felt emotions and emotions perceived in the artwork. We additionally highlight some other properties of emotions which we consider to be interesting from the perspective of robot art therapy: co-occurrence, referents, timing, and polysemy.

- *Co-occurrence.* Humans can feel multiple, sometimes opposite, emotions simultaneously [177,178]. For example, conflicting emotions might include enjoying a scary movie, a “dumb” joke, or cacophonous music; feeling happy and sad at experiencing a small winning or loss in gambling; feeling hopeful at new prospects but sad to lose contact with friends, when moving; feeling happy but sad when a child leaves home; or feeling happy for students who excel and sad for those who do not, after a test. Some examples of complimentary emotions are feeling relaxed and happy, or sad and angry. Humans appear to be adept at recognizing such emotions; it has even been suggested that blended emotions are displayed more often in the face than single emotions, and that people are better able to process this mixed information [179].
- *Referents.* Another property we note is that emotions are typically directed toward some context (here, a “referent”), and not random phenomena disconnected from the world; this can be seen in some typical patterns of thought occurring during the experiencing of emotions described by appraisal theory [180]. For instance, appraisal of an event can involve assessing its relevance, how difficult it is to handle, the causes, and norms for how people typically react [181]. Referents are also generally important for phenomena related to emotions such as “sentiments”, attitudes, preferences and opinions, which are typically directed toward something or someone. We believe that identifying the referent of an emotion is vital for human interactions; for example, a robot might need to know if a person is angry at the robot or at someone else. However, referents can be quite complex. A person cut off while driving might feel angry toward another driver, their car, or even everyone in the vicinity, or the situation in general. Moreover, in some cases, the referent might be obvious, but in others, even humans can have difficulty identifying why others feel the way they do, for example when someone is angry with them. Furthermore, it has been suggested that sometimes emotions are not directed at anything in particular, in the case of “core affect”, in which a person might feel excited, depressed, or relaxed [182].
- *Timing.* Additionally, emotions change over time. In general, there is an internal homeostatic tendency for strong emotions to gradually fade over time, and emotions can be more easily regulated as they start than when they are in progress [183]. The role of timing in “emotion episodes” has also been examined in affective events theory, in which it is claimed that that

the patterns with which emotions fluctuate over time are highly predictable [184]. In addition, some typical progressions of emotions have been described. For example, a grief response can proceed from denial, to anger, bargaining, depression, and finally acceptance [185]; in terms of basic emotions, this could be described as a progression from surprise to anger, sadness, and finally a neutral state. In addition, progressions of emotions from fear to anger are predicted in General Strain Theory [120,186]. We believe that humans have some intuitive understanding of such processes, which could be beneficial for a robot art therapist; for example, if it is known that scared people can become angry, predictions can be made about how a scared person might feel in the future.

- *Polysemy.* Additionally, emotional signals can be ambiguous. For example, nodding can be positive, indicating agreement, greeting, or thanks; neutral expressing confirmation; or negative conveying irony or emphatic insistence [124]. Robots should be aware of such nuances in order to communicate well with humans. In the current article, we seek to take into account such considerations, in exploring the complex phenomenon of the visual communication of emotions.

Appendix B. Creativity

Another core characteristic associated with art-making is creativity. Creativity has been defined in various ways: for example, by Cope as “the ability to associate two things which heretofore have not been considered particularly ... associate-able” [187] or as “imagination ... skill, and the ability to assess” [188]. A typical opinion reported by Arai and Lanier, among others, is that robots cannot be creative (<https://english.kyodonews.net/news/2018/07/528f2cc41122-feature-dont-fear-robots-fear-robotic-humans-says-japans-ai-auntie.html>), but can only recycle “data from people ... (where) the problem ... is that the people are made anonymous” [8]. This opinion is reflected in a statement by Jack Ma from Alibaba: “We have to teach (children) something unique, so a machine can never catch up with us ... we should teach our kids ... painting (and) art” (<https://www.youtube.com/watch?v=rHt-5-RyrJk>).

We agree that robots typically use data from people, but we believe this is not limited to computers, as people also draw insight from observing others; in addition, various machines described in Section 2 have already been developed to engage in basic art-making and painting. In the current article, we adopt a useful formulation from Winiger that creativity can be interpreted, not as something one has or doesn’t have, but rather as “a way of doing things” [8]. Moreover, we hold that creativity is not limited to some form such as poetry but appears in various different forms, in line with Horace’s “Quod libet audendi”, marked by some core component of novelty or incongruity; other factors such as usefulness are predictive of creativity only when the novelty is very high [189]. A creative artwork does not need to be novel for all humans, but can be creative for an individual or robot, as in Boden’s concept of P- versus H-creativity [190] or Simonton’s little-c versus Big-C creativity [191].

Appendix C. Basic Forms of Art Therapy

Although a comprehensive description is outside of the scope of the current article, we note a few examples of variants comprised within the umbrella term of “art therapy”: cognitive, behavioral, psychodynamic, and humanistic [71]. Cognitive approaches focus on mental processes as the key to therapy, whereas behavioral approaches focus on measurable outcomes. Psychodynamic approaches, such as Freudian or Jungian approaches, focus on unconscious dynamics and the phenomenon of transference. Furthermore, humanistic approaches, such as the person-centered approach, focus on a model of wellness rather than illness, and unconditional positive emotion and empathy toward the person making art.

Furthermore, for any particular kind of approach, numerous possibilities exist for how to seek to intervene to improve a person’s well-being. For example, Cognitive Behavior Therapy (CBT), combining ideas from cognitive and behavioral psychology, aims to confront and replace dysfunctional thought patterns and amend critical behaviors relevant to specific problems; high success rates

have resulted in CBT being suggested as a first step in the treatment of a majority of psychological problems [192]. Following such a strategy, a robot could use the A-B-C model prescribed in the first form of CBT, “Rational emotive behavior therapy”, to identify adversities (A), irrational beliefs (B)—such as that a person must always be successful, loved or comfortable—and consequences (C) [193]. Conversely, for difficult cases involving relapses or vague symptoms, schema therapy could be used to identify underlying rigid and maladaptive thought patterns, as well as emotional triggers (e.g., reflecting hypersensitivity to abandonment) [194]. Following the concept of guided imagery, the robot and human could jointly paint positive images, or the robot could paint positive images that the person could focus on, such as images depicting the overcoming of a problem [195]. Thought stopping could involve the robot detecting negative emotions projected into art and alerting the painter—although the efficacy of such techniques in isolation is debated [196]. Moreover, projective tests such as the Draw-A-Man Test, House–Tree–Person and Kinetic Family Drawing could be prescribed—although here too the reliability of such tests has been strongly questioned, due to the ambiguity which exists with the interpretation of symbols without feedback regarding the context, as well as confounds from artistic skills and culture [197].

Appendix D. The Emotional Meaning of Symbols

Some datasets of artwork exist, such as Painting-91, which comprises 4266 paintings from 91 different artists [65], but do not contain any information about emotions. However, some inferences can be drawn from image datasets assembled to investigate emotions [141,198,199]. In particular, the emotional meaning of some themes in images has been explored using the International Affective Picture System (IAPS), which contains roughly 1000 images rated by large numbers of people for valence and arousal [200]. Erotica, sports, adventure, and food were perceived as conveying high valence and arousal. Images of nature and babies were perceived to convey high valence and low arousal. Grieving scenes, illness and accidents, including a starving child and an injured face, were perceived to convey low valence and medium arousal. Furthermore, threatening content including snakes, guns, and violent death were perceived to convey low valence and high arousal. At the risk of over-simplification, these quadrants typically contain happy, relaxed, sad, and angry emotions, respectively. Note: Content perceived with low valence and arousal was rare, in line with Tellegen’s conjecture that low valence images are necessarily arousing. Strong positive correlations were also noted between how emotions were perceived by children and adults, as well as by women and men. We believe invaluable hints can be drawn from such results, but it is currently not possible to directly apply them to infer emotional meaning from symbols for robot art therapy. First, because the images typically contain multiple symbols expressing emotion, it is not straight-forward to infer from results for themes or pictures how individual symbols are perceived. For example, in one picture, which content is responsible for the valence and arousal levels attributed to the image: is it the expression or bent pose of a person, a caged dog beside him, or the natural landscape behind? Second, content can express meaning outside of basic sentiment; for example, it could be perceived as inappropriate if a robot sought to express happiness by painting an erotic scene for a person with dementia. Notwithstanding these concerns, we highlight one brief statement in this work, which was that over half of the dataset images depict people because it was claimed that the images which evoke the most emotion are those involving humans. Indeed, the images depicting sports and erotica, which are attributed joyful levels of valence and arousal, often contain people who appear joyful. Thus, we suggest that a robot could draw human faces with various facial expressions, as a basic symbol to express various emotions.

Another set of typical symbols which people use, from which inference could be possible, comprises emojis and emoticons. One study has started to investigate the emotional meaning of these common symbols [201]. Some drawbacks for use with robot art therapy however are that many of these symbols are used alongside text in a semantic rather than emotional way: for example, a picture of an airplane can refer to a trip, rather than joy or anger. This might explain some unintuitive results

in the above study such that symbols for a bomb, crying face, and worried face were labeled as slightly positive. Moreover, similar symbols were not grouped; there are many heart emojis, for example, which we expect would have similar emotional meanings. Furthermore, only valence/sentiment were measured, and not arousal or discrete emotions. Thus, we believe a gap exists in understanding the typical emotional meanings of symbols in paintings.

References

1. Gray, C. Art therapy: When pictures speak louder than words. *Can. Med. Assoc. J.* **1978**, *119*, 488. [PubMed]
2. Merback, M.B. *Perfection's Therapy: An Essay on Albrecht Dürer's Melencolia I*; MIT Press: Cambridge, MA, USA, 2018.
3. Toussaint, L.; Friedman, P. Forgiveness, Gratitude, and Well-Being: The Mediating Role of Affect and Beliefs. *J. Happiness Stud.* **2008**. [CrossRef]
4. Sperling, C.; Holst, K. Do Muddy Waters Shift Burdens. *Md. L. Rev.* **2016**, *76*, 629.
5. Jacobs, J. *Dynamic Drawing: Broadening Practice and Participation in Procedural Art*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2017.
6. Malchiodi, C. *The Art Therapy Sourcebook*; McGraw-Hill: New York, NY, USA, 2006.
7. Scherer, K.R. What are emotions? And how can they be measured? *Soc. Sci. Inf.* **2005**, *44*, 693–727. [CrossRef]
8. Gershgorn, D. Can we Make a Computer Make Art? In *Popular Science Special Edition: The New Artificial Intelligence*; Time Inc. Books: New York, NY, USA, 2016; pp. 64–67.
9. Holmén, K.; Ericsson, K.; Winblad, B. Social and emotional loneliness among non-demented and demented elderly people. *Arch. Gerontol. Geriatr.* **2000**, *31*, 177–192. [CrossRef]
10. Seppala, E.; King, M. Burnout at work isn't just about exhaustion. It's also about loneliness. *Harv. Bus. Rev.* **2017**, *29*. Available online: <https://hbr.org/2017/06/burnout-at-work-isnt-just-about-exhaustion-its-also-about-loneliness> (accessed on 30 August 2018).
11. Stuckey, H.L.; Nobel, J. The connection between art, healing, and public health: A review of current literature. *Am. J. Public Health* **2010**, *100*, 254–263. [CrossRef] [PubMed]
12. Rusted, J.; Sheppard, L.; Waller, D. A Multi-centre Randomized Control Group Trial on the Use of Art Therapy for Older People with Dementia. *Group Anal.* **2006**, *39*, 517–536. [CrossRef]
13. Bharatharaj, J.; Huang, L.; Mohan, R.E.; Al-Jumaily, A.; Krägeloh, C. Robot-assisted therapy for learning and social interaction of children with autism spectrum disorder. *Robotics* **2017**, *6*, 4. [CrossRef]
14. Wada, K.; Shibata, T. Robot therapy in a care house—its sociopsychological and physiological effects on the residents. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Orlando, FL, USA, 15–19 May 2006; pp. 3966–3971.
15. Wada, K.; Shibata, T. Living with seal robots—Its sociopsychological and physiological influences on the elderly at a care house. *IEEE Trans. Robot.* **2007**, *23*, 972–980. [CrossRef]
16. Bitonte, R.A.; De Santo, M. Art therapy: An underutilized, yet effective tool. *Ment. Illn.* **2014**, *6*, 5354. [CrossRef] [PubMed]
17. Evans, K.; Dubowski, J. *Art Therapy with Children on the Autistic Spectrum: Beyond Words*; Jessica Kingsley Publishers: London, UK, 2001.
18. Rivera, R.A. *Art Therapy for Individuals with Severe Mental Illness*. Master's Thesis, University of Southern California, Los Angeles, CA, USA, 2008.
19. Uttley, L.; Scope, A.; Stevenson, M.; Rawdin, A.; Buck, E.T.; Sutton, A.; Stevens, J.; Kaltenthaler, E.; Dent-Brown, K.; Wood, C. Systematic review and economic modelling of the clinical effectiveness and cost-effectiveness of art therapy among people with non-psychotic mental health disorders. *Health Technol. Assess.* **2015**, *19*, 1–120. [CrossRef] [PubMed]
20. Mihailidis, A.; Blunsden, S.; Boger, J.; Richards, B.; Zutis, K.; Young, L.; Hoey, J. Towards the development of a technology for art therapy and dementia: Definition of needs and design constraints. *Arts Psychother.* **2010**, *37*, 293–300. [CrossRef]
21. Kanamori, M.; Suzuki, M.; Oshiro, H.; Tanaka, M.; Inoguchi, T.; Takasugi, H.; Saito, Y.; Yokoyama, T. Pilot study on improvement of quality of life among elderly using a pet-type robot. In Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan, 16–20 July 2003.

22. Robinson, H.; MacDonald, B.; Kerse, N.; Broadbent, E. The Psychosocial Effects of a Companion Robot: A Randomized Controlled Trial. *J. Am. Med. Dir. Assoc.* **2013**, *14*, 661–667. [[CrossRef](#)] [[PubMed](#)]
23. Bartneck, C. Interacting with an embodied emotional character. In Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces, Pittsburgh, PA, USA, 23–26 June 2003; pp. 55–60, doi:10.1145/782896.782911. [[CrossRef](#)]
24. Leyzberg, D.; Spaulding, S.; Toneva, M.; Scassellati, B. The Physical Presence of a Robot Tutor Increases Cognitive Learning Gains. In Proceedings of the Annual Meeting of the Cognitive Science Society, Sapporo, Japan, 1–4 August 2012.
25. Powers, A.; Kiesler, S.; Fussell, S.; Torrey, C. Comparing a Computer Agent with a Humanoid Robot. In Proceedings of the 2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI'07), Arlington, VA, USA, 9–11 March 2007.
26. Shinozawa, K.; Naya, F.; Yamato, J.; Kogure, K. Differences in effect of robot and screen agent recommendations on human decision-making. *Int. J. Hum. Comput. Stud.* **2005**, *62*, 267–279. [[CrossRef](#)]
27. Cave, S.; SOhÉigeartaigh, S. An AI Race for Strategic Advantage: Rhetoric and Risks. In Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics and Society, New Orleans, LA, USA, 2–3 February 2018.
28. Colton, S.; Wiggins, G.A. Computational Creativity: The Final Frontier? In Proceedings of the 20th European Conference on Artificial Intelligence ECAI, Montpellier, France, 27–31 August 2012; Volume 12, pp. 21–26.
29. Picard, R.W. *Affective Computing*; M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321; MIT Press: Cambridge, MA, USA, 1995.
30. Wang, Z.; Xie, L.; Lu, T. Research progress of artificial psychology and artificial emotion in China. *CAAI Trans. Intell. Technol.* **2016**, *1*, 355–365. [[CrossRef](#)]
31. Cooney, M.; Pashami, S.; Sant'Anna, A.; Fan, Y.; Nowaczyk, S. Pitfalls of Affective Computing: How can the automatic visual communication of emotions lead to harm, and what can be done to mitigate such risks? In Proceedings of the WWW'18 Companion: The 2018 Web Conference Companion, Lyon, France, 23–27 April 2018; ACM: New York, NY, USA, 2018; doi:10.1145/3184558.3191611.
32. Narasimman, S.V.; Westerlund, D. Robot Artwork Using Emotion Recognition. Master's Thesis, Halmstad University, Halmstad, Sweden, 2017.
33. Badnjević, A.; Gurbeta, L. Development and perspectives of biomedical engineering in South East European countries. In Proceedings of the 39th IEEE International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 30 May–3 June 2016; pp. 457–460.
34. Guan, X.; Ji, L.; Wang, R. Development of Exoskeletons and Applications on Rehabilitation. In Proceedings of the 2015 International Conference on Mechanical Engineering and Electrical Systems (ICMES 2015), Singapore, 16–18 December 2015; EDP Sciences: Les Ulis, France, 2016; Volume 40; p. 02004.
35. Fasola, J.; Mataric, M.J. Using socially assistive human–robot interaction to motivate physical exercise for older adults. *Proc. IEEE* **2012**, *100*, 2512–2526. [[CrossRef](#)]
36. Hansen, S.T.; Bak, T.; Risager, C. An adaptive game algorithm for an autonomous, mobile robot—A real world study with elderly users. In Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication, Paris, France, 9–13 September 2012; pp. 892–897.
37. Hebesberger, D.V.; Dondrup, C.; Gisinger, C.; Hanheide, M. Patterns of use: How older adults with progressed dementia interact with a robot. In Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, Vienna, Austria, 6–9 March 2017; pp. 131–132.
38. Moyle, W.; Jones, C.; Sung, B.; Bramble, M.; O'Dwyer, S.; Blumenstein, M.; Estivill-Castro, V. What effect does an animal robot called CuDDler have on the engagement and emotional response of older people with dementia? A pilot feasibility study. *Int. J. Soc. Robot.* **2016**, *8*, 145–156. [[CrossRef](#)]
39. Plaisant, C.; Druin, A.; Lathan, C.; Dakhane, K.; Edwards, K.; Vice, J.M.; Montemayor, J. A storytelling robot for pediatric rehabilitation. In Proceedings of the Fourth International ACM Conference on Assistive Technologies, Arlington, VA, USA, 13–15 November 2000; pp. 50–55.
40. Dehkordi, P.S.; Moradi, H.; Mahmoudi, M.; Pouretamad, H.R. The design, development, and deployment of RoboParrot for screening autistic children. *Int. J. Soc. Robot.* **2015**, *7*, 513–522. [[CrossRef](#)]
41. Gonzalez-Pacheco, V.; Ramey, A.; Alonso-Martín, F.; Castro-Gonzalez, A.; Salichs, M.A. Maggie: A social robot as a gaming platform. *Int. J. Soc. Robot.* **2011**, *3*, 371–381. [[CrossRef](#)]

42. Pop, C.A.; Simut, R.; Pinteau, S.; Saldien, J.; Rusu, A.; David, D.; Vanderfaellie, J.; Lefeber, D.; Vanderborght, B. Can the social robot Probo help children with autism to identify situation-based emotions? A series of single case experiments. *Int. J. Hum. Robot.* **2013**, *10*, 1350025. [[CrossRef](#)]
43. Salvador, M.J.; Silver, S.; Mahoor, M.H. An emotion recognition comparative study of autistic and typically-developing children using the zeno robot. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015; pp. 6128–6133.
44. Greczek, J.; Kaszubski, E.; Atrash, A.; Matarić, M. Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders. In Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Edinburgh, UK, 25–29 August 2014; pp. 561–566.
45. Srinivasan, S.M.; Kaur, M.; Park, I.K.; Gifford, T.D.; Marsh, K.L.; Bhat, A.N. The effects of rhythm and robotic interventions on the imitation/praxis, interpersonal synchrony, and motor performance of children with autism spectrum disorder (ASD): A pilot randomized controlled trial. *Autism Res. Treat.* **2015**, *2015*, 736516. [[CrossRef](#)] [[PubMed](#)]
46. Kosuge, K.; Hayashi, T.; Hirata, Y.; Tobiyama, R. Dance Partner Robot-Ms DanceR. In Proceedings of the IROS 2003, Las Vegas, NV, USA, 27–31 October 2003.
47. Kozima, H.; Michalowski, M.P.; Nakagawa, C. Keepon. *Int. J. Soc. Robot.* **2009**, *1*, 3–18. [[CrossRef](#)]
48. Tapus, A.; Mataric, M.J. Socially assistive robotic music therapist for maintaining attention of older adults with cognitive impairments. In Proceedings of the AAAI Fall Symposium AI in Eldercare: New Solutions to Old Problems, Washington, DC, USA, 7–9 November 2008.
49. Lim, A.; Mizumoto, T.; Cahier, L.K.; Otsuka, T.; Takahashi, T.; Komatani, K.; Ogata, T.; Okuno, H.G. Robot musical accompaniment: Integrating audio and visual cues for real-time synchronization with a human flutist. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Taipei, Taiwan, 18–22 October 2010; pp. 1964–1969.
50. Martín, F.; Agüero, C.E.; Cañas, J.M.; Valenti, M.; Martínez-Martín, P. Robototherapy with dementia patients. *Int. J. Adv. Robot. Syst.* **2013**, *10*, 10. [[CrossRef](#)]
51. Graf, B.; Reiser, U.; Hägele, M.; Mauz, K.; Klein, P. Robotic Home Assistant Care-O-bot[®]3-Product Vision and Innovation Platform. In Proceedings of the 2009 IEEE Workshop on Advanced Robotics and Its Social Impacts (ARSO 2009), Tokyo, Japan, 23–25 November 2009; pp. 139–144.
52. Odashima, T.; Onishi, M.; Tahara, K.; Takagi, K.; Asano, F.; Kato, Y.; Nakashima, H.; Kobayashi, Y.; Mukai, T.; Luo, Z.; et al. A Soft Human-Interactive Robot RIMAN. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Beijing, China, 9–15 October 2006.
53. Pollack, M.; Engberg, S.; Thrun, S.; Brown, L.; Colbry, D.; Orosz, C.; Peintner, B.; Ramakrishnan, S.; Dunbar-Jacob, J.; McCarthy, C. Pearl: A Mobile Robotic Assistant for the Elderly. In Proceedings of the AAAI Workshop on Automation as Caregiver, Edmonton, AB, Canada, 28–29 July 2002.
54. Richardson, K.; Coeckelbergh, M.; Wakunuma, K.; Billing, E.; Ziemke, T.; Gomez, P.; Vanderborght, B.; Belpaeme, T. Robot Enhanced Therapy for Children with Autism (DREAM): A Social Model of Autism. *IEEE Technol. Soc. Mag.* **2018**, *37*, 30–39. [[CrossRef](#)]
55. McClelland, R.T. Robotic Alloparenting: A New Solution to an Old Problem? *J. Mind Behav.* **2016**, *37*.
56. Liu, X.; Wu, Q.; Zhao, W.; Luo, X. Technology-Facilitated Diagnosis and Treatment of Individuals with Autism Spectrum Disorder: An Engineering Perspective. *Appl. Sci.* **2017**, *7*, 1051. [[CrossRef](#)]
57. Mordoch, E.; Osterreicher, A.; Guse, L.; Roger, K.; Thompson, G. Use of social commitment robots in the care of elderly people with dementia: A literature review. *Maturitas* **2013**, *74*, 14–20. [[CrossRef](#)] [[PubMed](#)]
58. Edwards, B. The Never-Before-Told Story of the World’s First Computer Art (It’s a Sexy Dame). *The Atlantic*, 24 January 2013.
59. Brown, P. From systems art to artificial life: Early generative art at the Slade School of Fine Art. In *White Heat Cold Logic: British Computer Art*; MIT Press: Cambridge, CA, USA, 2008; pp. 275–290.
60. Srikaew, A.; Cambron, M.E.; Northrup, S.; Peters, R.A., II; Wilkes, M.; Kawamura, K. Humanoid drawing robot. In Proceedings of the IASTED International Conference on Robotics and Manufacturing, Banff, AB, Canada, 1–4 July 1998.
61. Kudoh, S.; Ogawara, K.; Ruchanurucks, M.; Ikeuchi, K. Painting robot with multi-fingered hands and stereo vision. *Robot. Auton. Syst.* **2009**, *57*, 279–288. [[CrossRef](#)]

62. Calinon, S.; Epiney, J.; Billard, A. A humanoid robot drawing human portraits. In Proceedings of the 2005 5th IEEE RAS International Conference on Humanoid Robots, Tsukuba, Japan, 5 December 2005; pp. 161–166.
63. Lee, Y.J.; Zitnick, C.L.; Cohen, M.F. Shadowdraw: Real-time user guidance for freehand drawing. In Proceedings of the ACM SIGGRAPH 2011 Papers (SIGGRAPH'11), Vancouver, BC, Canada, 7–11 August 2011; pp. 27:1–27:10.
64. Ha, D.; Eck, D. A neural representation of sketch drawings. *arXiv* **2017**, arXiv:1704.03477.
65. Khan, F.S.; Beigpour, S.; van de Weijer, J.; Felsberg, M. Painting-91: A large scale database for computational painting categorization. *Mach. Vis. Appl.* **2014**, *25*, 1385–1397. [[CrossRef](#)]
66. Crick, C.; Scassellati, B. Inferring narrative and intention from playground games. In Proceedings of the 12th IEEE Conference on Development and Learning, Monterey, CA, USA, 9–12 August 2008.
67. Huang, C.; Mutlu, B. Anticipatory Robot Control for Efficient Human-Robot Collaboration. In Proceedings of the ACM/IEEE International Conference on Human Robot Interaction, Christchurch, New Zealand, 7–10 March 2016; IEEE Press: Piscataway, NJ, USA, 2016; pp. 83–90.
68. Montebelli, A.; Tykal, M. Intention Disambiguation: When does action reveal its underlying intention? In Proceedings of the HRI 2017, Vienna, Austria, 6–9 March 2017.
69. Breazeal, C.; Buchsbaum, D.; Gray, J.; Gatenby, D.; Blumberg, B. Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artif. Life* **2005**, *11*, 31–62. [[CrossRef](#)] [[PubMed](#)]
70. Broekens, J. Emotion and reinforcement: Affective facial expressions facilitate robot learning. In *Artificial Intelligence for Human Computing, LNAI 4451*; Huang, T.S., Nijholt, A., Pantic, M., Pentland, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 113–132.
71. Rubin, J.A. *Art Therapy: An Introduction*; Psychology Press: Wellington, New Zealand, 1999.
72. Hogan, S. (Ed.) *Feminist Approaches to Art Therapy*; Psychology Press: Wellington, New Zealand, 1997.
73. Talwar, S.; Iyer, J.; Doby-Copeland, C. The invisible veil: Changing paradigms in the art therapy profession. *Art Ther.* **2004**, *21*, 44–48. [[CrossRef](#)]
74. Tapus, A.; Mataric, M.J. Emulating Empathy in Socially Assistive Robotics. In Proceedings of the AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics, Palo Alto, CA, USA, 26–28 March 2007; pp. 93–96.
75. Kahn, P.H., Jr.; Ishiguro, H.; Friedman, B.; Kanda, T.; Freier, N.G.; Severson, R.L.; Miller, J. What is a Human? Toward psychological benchmarks in the field of human–robot interaction. *Interact. Stud.* **2007**, *8*, 363–390.
76. Hart, G.J. The five W's: An old tool for the new task of audience analysis. *Tech. Commun.* **1996**, *43*, 139–145.
77. Kugel, P. How Professors Develop as Teachers. *Stud. High. Educ.* **1993**, *18*, 315–328. [[CrossRef](#)]
78. Mills, A. The diagnostic drawing series. In *Handbook of Art Therapy*; Guilford Press: New York, NY, USA, 2003; pp. 401–409.
79. Muri, S.A. Beyond the face: Art therapy and self-portraiture. *Arts Psychother.* **2007**, *34*, 331–339. [[CrossRef](#)]
80. Riether, N.; Hegel, F.; Wrede, B.; Horstmann, G. Social facilitation with social robots? In Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction, Boston, MA, USA, 5–8 March 2012; pp. 41–48.
81. Killeen, J.P.; Evans, G.W.; Danko, S. The role of permanent student artwork in students' sense of ownership in an elementary school. *Environ. Behav.* **2003**, *35*, 250–263. [[CrossRef](#)]
82. Turk, D.J.; van Bussel, K.; Waiter, G.; Macrae, C.N. Mine and me: Exploring the neural basis of object ownership. *J. Cognit. Neurosci.* **2011**, *23*, 3657–3668. [[CrossRef](#)] [[PubMed](#)]
83. Phillips, J. Working with adolescents' violent imagery. In *Handbook of Art Therapy*; Guilford Press: New York, NY, USA, 2003; pp. 229–238.
84. Goetz, J.; Kiesler, S.; Powers, A. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication, Millbrae, CA, USA, 2 November 2003; pp. 55–60.
85. Tapus, A.; Țăpuș, C.; Matarić, M.J. User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intell. Serv. Robot.* **2008**, *1*, 169. [[CrossRef](#)]
86. Gross, J.J. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* **2002**, *39*, 281–291. [[CrossRef](#)] [[PubMed](#)]
87. Byrne, D.; Clore, G.L.; Smeaton, G. The attraction hypothesis: Do similar attitudes affect anything? *J. Personal. Soc. Psychol.* **1986**, *51*, 1167–1170. [[CrossRef](#)]

88. Hanson, D. Exploring the aesthetic range for humanoid robots. In Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, Vancouver, BC, Canada, 26–29 July 2006; pp. 39–42.
89. Mara, M.; Appel, M. Science fiction reduces the eeriness of android robots: A field experiment. *Comput. Hum. Behav.* **2015**, *48*, 156–162. [[CrossRef](#)]
90. Campos, J.J. When the negative becomes positive and the reverse: Comments on Lazarus’s critique of positive psychology. *Psychol. Inq.* **2003**, *14*, 110–113.
91. Drake, J.E.; Winner, E. Confronting Sadness Through Art-Making: Distraction Is More Beneficial than Venting. *Psychol. Aesthet. Creat. Arts* **2012**, *6*, 255–261. [[CrossRef](#)]
92. Conrad, P. It’s boring: Notes on the meanings of boredom in everyday life. *Qual. Sociol.* **1997**, *20*, 465–475. [[CrossRef](#)]
93. Cooney, M.; Nishio, S.; Ishiguro, H. Affectionate Interaction with a Small Humanoid Robot Capable of Recognizing Social Touch Behavior. *ACM Trans. Interact. Intell. Syst.* **2014**, *4*, 1–32. [[CrossRef](#)]
94. Hobbes, T. Human Nature. In *The English Works of Thomas Hobbes of Malmesbury*; Molesworth, W., Ed.; Bohn: London, UK, 1840; Volume IV.
95. Kim, S.; Bak, J.; Oh, A.H. Do You Feel What I Feel? Social Aspects of Emotions in Twitter Conversations. In Proceedings of the ICWSM, Dublin, Ireland, 4–7 June 2012; AAAI Press: Palo Alto, CA, USA.
96. Giles, H.; Baker, S.C. Communication accommodation theory. *Int. Encycl. Commun.* **2008**, doi:10.1002/9781405186407.wbiecc067. [[CrossRef](#)]
97. Damiano, L.; Dumouchel, P.; Lehmann, H. Artificial Empathy: An Interdisciplinary Investigation. *Int. J. Soc. Robot.* **2015**, *7*, 3–5. [[CrossRef](#)]
98. De Vignemont, F.; Singer, T. The empathic brain: How, when and why? *Trends Cognit. Sci.* **2006**, *10*, 435–441. [[CrossRef](#)] [[PubMed](#)]
99. Rodrigues, S.H.; Mascarenhas, S.; Dias, J.; Paiva, A. A process model of empathy for virtual agents. *Interact. Comput.* **2014**, *27*, 371–391. [[CrossRef](#)]
100. Alenljung, B.; Andreasson, R.; Billing, E.A.; Lindblom, J.; Lowe, R. User Experience of Conveying Emotions by Touch. In Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Lisbon, Portugal, 28 August–1 September 2017; pp. 1240–1247, doi:10.1109/ROMAN.2017.8172463. [[CrossRef](#)]
101. Mehrabian, A.; Ferris, S.R. Inference of attitudes from nonverbal communication in two channels. *J. Consult. Psychol.* **1967**, *31*, 3. [[CrossRef](#)]
102. Kobayashi, K.; Kitamura, Y.; Yamada, S. Action sloping as a way for users to notice a robot’s function. In Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN’07), Jeju, Korea, 26–29 August 2007; pp. 445–450; doi:10.1109/ROMAN.2007.4415125. [[CrossRef](#)]
103. Christiansen, L.H.; Frederiksen, N.Y.; Jensen, B.S.; Ranch, A.; Skov, M.B.; Thiruravichandran, N. Don’t look at me, I’m talking to you: Investigating input and output modalities for in-vehicle systems. In Proceedings of the IFIP Conference on Human-Computer Interaction, Lisbon, Portugal, 5–9 September 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 675–691.
104. Frederick, S.; Loewenstein, G. 16 Hedonic Adaptation. In *Well-Being. The Foundations of Hedonic Psychology*; Kahneman, D., Diener, E., Schwarz, N., Eds.; Russell Sage: New York, NY, USA, 1999; pp. 302–329.
105. Cooney, M.; Nishio, S.; Ishiguro, H. Designing Robots for Well-being: Theoretical Background and Visual Scenes of Affectionate Play with a Small Humanoid Robot. *Lovotics* **2014**, *1*, 101. [[CrossRef](#)]
106. Cooney, M.; Sant’Anna, A. Avoiding Playfulness Gone Wrong: Exploring Multi-objective Reaching Motion Generation in a Social Robot. *Int. J. Soc. Robot.* **2017**. [[CrossRef](#)]
107. Ryff, C.D. Happiness Is Everything, or Is It? Explorations on the Meaning of Psychological Well-Being. *J. Personal. Soc. Psychol.* **1989**, *57*, 1069–1081. [[CrossRef](#)]
108. Fava, G.A.; Ruini, C. Development and characteristics of a well-being enhancing psychotherapeutic strategy: Well-being therapy. *J. Behav. Ther. Exp. Psychiatry* **2003**, *34*, 45–63. [[CrossRef](#)]
109. Korb, A. *The Upward Spiral: Using Neuroscience to Reverse the Course of Depression, One Small Change at a Time*; New Harbinger Publications: Oakland, CA, USA, 2015.
110. Williams, L.E.; Bargh, J.A. Experiencing physical warmth promotes interpersonal warmth. *Science* **2008**, *322*, 606–607. [[CrossRef](#)] [[PubMed](#)]

111. Jung, M.M.; Poel, M.; Poppe, R.; Heylen, D.K. Automatic recognition of touch gestures in the corpus of social touch. *J. Multimodal User Interfaces* **2017**, *11*, 81–96. [[CrossRef](#)]
112. Shiomi, M.; Nakagawa, K.; Shinozawa, K.; Matsumura, R.; Ishiguro, H.; Hagita, N. Does a robot's touch encourage human effort? *Int. J. Soc. Robot.* **2017**, *9*, 5–15. [[CrossRef](#)]
113. Silvera-Tawil, D.; Rye, D.; Velonaki, M. Artificial skin and tactile sensing for socially interactive robots: A review. *Robot. Auton. Syst.* **2015**, *63*, 230–243. [[CrossRef](#)]
114. Van Erp, J.B.; Toet, A. Social touch in human–computer interaction. *Front. Digit. Hum.* **2015**, *2*, 2.
115. Nomura, T.; Uratani, T.; Matsumoto, K.; Kanda, T.; Kidokoro, H.; Suehiro, Y.; Yamada, S. Why do children abuse robots? In Proceedings of the International Conference on Human–Robot Interaction'15, Portland, Oregon, 2–5 March 2015.
116. Cooper, A.J.; Mendonca, J.D. A prospective study of patient assaults on nurses in a provincial psychiatric hospital in Canada. *Acta Psychiatr. Scand.* **1991**, *84*, 163–166. [[CrossRef](#)] [[PubMed](#)]
117. Bartneck, C.; Nomura, T.; Kanda, T.; Suzuki, T.; Kennsuke, K. A cross-cultural study on attitudes towards robots. In Proceedings of the HCI International, Las Vegas, NV, USA, 22–27 July 2005.
118. Corrigan, P. How stigma interferes with mental health care. *Am. Psychol.* **2004**, *59*, 614. [[CrossRef](#)] [[PubMed](#)]
119. Turkle, S. *Alone Together: Why We Expect more from Technology and less from Each Other*; Basic Books; Hachette UK: London, UK, 2011.
120. Agnew, R. General strain theory: Current status and directions for further research. *Tak. Stock* **2006**, *15*, 101–123.
121. Haslam, N. Dehumanization: An integrative review. *Personal. Soc. Psychol. Rev.* **2006**, *10*, 252–264. [[CrossRef](#)] [[PubMed](#)]
122. Scharer, E. Media exposure and sensitivity to violence in news reports: Evidence of desensitization? *J. Mass Commun. Q.* **2008**, *85*, 291–310. [[CrossRef](#)]
123. Ekman, P.; Friesen, W.V. Nonverbal leakage and clues to deception. *Psychiatry* **1969**, *32*, 88–106. [[CrossRef](#)] [[PubMed](#)]
124. Poggi, I.; D'Errico, F.; Vincze, L. Types of Nods. The Polysemy of a Social Signal. In Proceedings of the LREC 2010, Valetta, Malta, 19–21 May 2010.
125. O'Neill, C. *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*; Crown, Penguin Random House LLC: New York, NY, USA, 2016.
126. Bullington, J. 'Affective' Computing and Emotion Recognition Systems: The Future of Biometric Surveillance? In Proceedings of the 2nd Annual Conference on Information Security Curriculum Development (InfoSecCD'05), Kennesaw, Georgia, 23–24 September 2008; ACM: New York, NY, USA, 2005; pp. 95–99, doi:10.1145/1107622.1107644.
127. Grote, T.; Korn, O. Risks and Potentials of Affective Computing. An Interdisciplinary View on the ACM Code of Ethics. In Proceedings of the CHI 2017 Workshop on Ethical Encounters in HCI, Denver, CO, USA, 6–11 May 2017.
128. Sharkey, A.; Wood, N. The Paro seal robot: Demeaning or enabling. In Proceedings of the AISB, London, UK, 1–4 April 2014; Volume 36.
129. Ullman, D.; Leite, I.; Phillips, J.; Kim-Cohen, J.; Scassellati, B. Smart human, smarter robot: How cheating affects perceptions of social agency. In Proceedings of the 36th Annual Conference of the Cognitive Science Society (CogSci 2014), Quebec City, QC, Canada, 23–26 July 2014.
130. Chadalavada, R.T.; Andreasson, H.; Krug, R.; Lilienthal, A.J. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In Proceedings of the IEEE 2015 European Conference on Mobile Robots (ECMR), Lincoln, UK, 2–4 September 2015; pp. 1–6.
131. Plutchik, R. Emotions: A general psychoevolutionary theory. In *Approaches to Emotion*, Scherer, K.R., Ekman, P., Eds; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 1984; pp. 197–219.
132. Ståhl, A.; Sundström, P.; Höök, K. A foundation for emotional expressivity. In Proceedings of the 2005 Conference on Designing for User Experience, San Francisco, CA, USA, 3–5 November 2005; American Institute of Graphic Arts: New York, NY, USA, 2005; p. 33.
133. Meier, B.P.; Robinson, M.D.; Clore, G.L. Why good guys wear white automatic inferences about stimulus valence based on brightness. *Psychol. Sci.* **2004**, *15*, 82–87. [[CrossRef](#)] [[PubMed](#)]
134. Bertamini, M.; Palumbo, L.; Gheorghes, T.N.; Galatsidas, M. Do observers like curvature or do they dislike angularity? *Br. J. Psychol.* **2016**, *107*, 154–178. [[CrossRef](#)] [[PubMed](#)]

135. Enquist, M.; Arak, A. Symmetry, beauty, and evolution. *Nature* **1994**, *372*, 169–172. [[CrossRef](#)] [[PubMed](#)]
136. Grammer, K.; Thornhill, R. Human facial attractiveness and sexual selection: The role of symmetry and averageness. *J. Comp. Psychol.* **1994**, *108*, 233–242. [[CrossRef](#)] [[PubMed](#)]
137. Joshi, D.; Datta, R.; Fedorovskaya, E.; Luong, Q.; Wang, J.Z.; Li, J.; Luo, J. Aesthetics and emotions in images. *IEEE Signal Process. Mag.* **2011**, *28*, 94–115. [[CrossRef](#)]
138. Palmer, S.E.; Schloss, K.B.; Sammartino, J. Visual aesthetics and human preference. *Annu. Rev. Psychol.* **2013**, *64*, 77–107. [[CrossRef](#)] [[PubMed](#)]
139. Lauer, D.A.; Pentak, S. *Design Basics*; Cengage Learning: Boston, MA, USA, 2011.
140. Rodin, R. Mood Lines: Setting the Tone of Your Design. 2015. Available online: <https://zevendesign.com/mood-lines-giving-designs-attitude/> (accessed on 15 June 2018).
141. Machajdik, J.; Hanbury, A. Affective Image Classification using Features Inspired by Psychology and Art Theory. In Proceedings of the 18th ACM international conference on Multimedia (MM'10), Firenze, Italy, 25–29 October 2010.
142. Shechtman, E.; Irani, M. Matching local self-similarities across images and videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.
143. Crowley, E.; Zisserman, A. The State of the Art: Object Retrieval in Paintings using Discriminative Regions. In Proceedings of the British Machine Vision Conference, Nottingham, UK, 1–5 September 2014.
144. Jongejan, J.; Rowley, H.; Kawashima, T.; Kim, J.; Fox-Gieg, N. The Quick, Draw!—A.I. Experiment. 2016. Available online: <https://quickdraw.withgoogle.com/> (accessed on 30 August 2018).
145. Motzenbecker, D. Fast Drawing for Everyone. 2017. Available online: <https://www.blog.google/technology/ai/fast-drawing-everyone/> (accessed on 30 August 2018).
146. Ng, H.; Nguyen, V.D.; Vonikakis, V.; Winkler, S. Deep Learning for Emotion Recognition on Small Datasets Using Transfer Learning. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI'15), Seattle, WA, USA, 9–13 November 2015; ACM: New York, NY, USA, 2015; pp. 443–449. [[CrossRef](#)]
147. Picard, R.W.; Healey, J. Affective wearables. *Pers. Technol.* **1997**, *1*, 231–240. [[CrossRef](#)]
148. Samara, A.; Menezes, M.L.R.; Galway, L. Feature Extraction for Emotion Recognition and Modelling Using Neurophysiological Data. In Proceedings of the International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), Granada, Spain, 14–16 December 2016; pp. 138–144.
149. Leviathan, Y.; Matias, Y. Google Duplex: An AI System for Accomplishing Real World Tasks Over the Phone. Google AI Blog, 2018. Available online: <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html> (accessed on 30 August 2018).
150. Cambria, E.; Poria, S.; Gelbukh, A.; Thelwall, M. Sentiment analysis is a big suitcase. *IEEE Intell. Syst.* **2017**, *32*, 74–80. [[CrossRef](#)]
151. Breazeal, C.; Aryananda, L. Recognition of affective communicative intent in robot-directed speech. *Auton. Robots* **2002**, *12*, 83–104. [[CrossRef](#)]
152. Gold, K.; Doniec, M.; Crick, C.; Scassellati, B. Robotic vocabulary building using extension inference and implicit contrast. *Artif. Intell.* **2009**, *173*, 145–166. [[CrossRef](#)]
153. Rani, P.; Sarkar, N.; Smith, C.; Kirby, L. Anxiety detecting robotic system—Towards implicit human–robot collaboration. *Robotica* **2004**, *22*, 85–95. [[CrossRef](#)]
154. Xu, H.; Plataniotis, K.N. Affect recognition using EEG signal. In Proceedings of the 14th IEEE International Workshop on Multimedia Signal Processing (MMSp), Banff, AB, Canada, 17–19 September 2012; pp. 299–304.
155. Merla, A.; Romani, G.L. Thermal signatures of emotional arousal: A functional infrared imaging study. In Proceedings of the 29th Annual IEEE International Conference on Engineering in Medicine and Biology Society (EMBS 2007), Lyon, France, 22–26 August 2007; pp. 247–249.
156. Liu, Z.; Wang, S. Emotion recognition using hidden markov models from facial temperature sequence. In *Affective Computing and Intelligent Interaction*; Springer: Berlin, Germany, 2011; pp. 240–247.
157. Hadjeres, G.; Pachet, F.; Nielsen, F. DeepBach: A Steerable Model for Bach Chorales Generation. *ArXiv* **2016**, arXiv:1612.01010.
158. Lebrun, T. Who Is the Artificial Author? In Proceedings of the Canadian Conference on Artificial Intelligence, Edmonton, AB, Canada, 16–19 May 2017; Springer: Cham, Switzerland, 2017; pp. 411–415.

159. Newitz, A. Movie Written by Algorithm Turns out to Be Hilarious and Intense. *Ars Technica*, 2016. Available online: <https://arstechnica.com/gaming/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/> (accessed on 25 June 2018).
160. Cook, M.; Colton, S.; Gow, J. The ANGELINA Videogame Design System—Part I. *IEEE Trans. Comput. Intell. AI Games* **2017**, *9*, 192–203. [CrossRef]
161. Netzer, Y.; Gabay, D.; Goldberg, Y.; Elhadad, M. Gaiku: Generating Haiku with word associations norms. In Proceedings of the Workshop on Computational Approaches to Linguistic Creativity (CALC'09), Boulder, CO, USA, 4 June 2009; Association for Computational Linguistics: Stroudsburg, PA, USA, 2009; pp. 32–39.
162. Elgammal, A.; Liu, B.; Elhoseiny, M.; Mazzone, M. CAN: Creative Adversarial Networks, Generating Art by Learning About Styles and Deviating from Style Norms. *arXiv* **2017**, arXiv:1706.07068.
163. Kauffman, S.A. *Investigations*; Oxford University Press: Oxford, UK, 2000.
164. Loreto, V.; Servedio, V.D.; Strogatz, S.H.; Tria, F. Dynamics on expanding spaces: Modeling the emergence of novelties. In *Creativity and Universality in Language*; Springer: Cham, Switzerland, 2016; pp. 59–83.
165. Gunning, D. Explainable Artificial Intelligence (XAI). Technical Report by the Defense Advanced Research Projects Agency (DARPA). 2017. Available online: <http://www.darpa.mil/attachments/DARPA-BAA-16-53.pdf> (accessed on 27 June 2018).
166. Schindler, M.; Lilienthal, A.J.; Chadalavada, R.; Ögren, M. Creativity in the eye of the student. Refining investigations of mathematical creativity using eye-tracking goggles. In Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education (PME 40), Szeged, Hungary, 3–7 August 2016.
167. Fox, M.; Long, D.; Magazzeni, D. Explainable Planning. In Proceedings of the IJCAI-17 Workshop on Explainable AI (XAI), Melbourne, Australia, 20 August 2017.
168. Engelberg, D.; Seffah, A. A Framework for Rapid Mid-Fidelity Prototyping of Web Sites. In *Usability: Gaining a Competitive Edge, Proceedings of the IFIP World Computer Congress, Deventer, The Netherlands, 25–30 August 2002*; Hammond, J., Gross, T., Wesson, J., Eds.; Springer: Boston, MA, USA, 2002.
169. McClellan, J.H.; Parks, T.W. A unified approach to the design of optimum FIR linear phase digital filters. *IEEE Trans. Circuit Theory* **1973**, *CT-20*, 697–701. [CrossRef]
170. Welch, P. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **1967**, *15*, 70–73. [CrossRef]
171. Lovett, A.; Scassellati, B. Using a robot to reexamine looking time experiments. In Proceedings of the Third International Conference on Development and Learning, La Jolla, CA, USA, 20–22 October 2004; pp. 284–291.
172. Darwin, C.R. *The Expression of the Emotions in Man and Animals*, 1st ed.; John Murray: London, UK, 1872. Available online: <http://darwin-online.org.uk/content/frameset?itemID=F1142&viewtype=text&pageseq=1> (accessed on 4 November 2014).
173. Ekman, P. An argument for basic emotions. *Cognit. Emot.* **1992**, *6*, 169–200. [CrossRef]
174. Russell, J.A. A circumplex model of affect. *J. Personal. Soc. Psychol.* **1980**, *39*, 1161–1178. [CrossRef]
175. Wundt, W. *Outlines of Psychology*; Wilhelm Englemann: Leipzig, Germany, 1897.
176. Soleymani, M.; Garcia, D.; Jou, B.; Schuller, B.; Chang, S.; Pantic, M. A survey of multimodal sentiment analysis. *Image Vis. Comput.* **2017**, *65*, 3–14. [CrossRef]
177. Cacioppo, J.T.; Gardner, W.L.; Berntson, G.G. The Affect System Has Parallel and Integrative Processing Components: Form Follows Function. *J. Personal. Soc. Psychol.* **1999**, *76*, 839–855. [CrossRef]
178. Hong, J.; Lee, A.Y. Feeling Mixed but Not Torn: The Moderating Role of Construal Level in Mixed Emotions Appeals. *J. Consum. Res.* **2010**, *37*. [CrossRef]
179. LaPlante, D.; Ambady, N. Multiple messages: Facial recognition advantage for compound expressions. *J. Nonverbal Behav.* **2000**, *24*, 211–224. [CrossRef]
180. Ellsworth, P.C.; Scherer, K.R. Appraisal processes in emotion. In *Series in affective science. Handbook of Affective Sciences*; Davidson, R.J., Scherer, K.R., Goldsmith, H.H., Eds.; Oxford University Press: New York, NY, USA, 2007; pp. 572–595.
181. Scherer, K.R.; Schorr, A.; Johnstone, T. (Eds.) *Appraisal Processes in Emotion: Theory, Methods, Research*; Oxford University Press: Oxford, UK, 2001.
182. Russell, J.A.; Barrett, L.F. Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *J. Personal. Soc. Psychol.* **1999**, *76*, 805. [CrossRef]

183. Sheppes, G.; Gross, J.J. Is timing everything? Temporal considerations in emotion regulation. *Personal. Soc. Psychol. Rev.* **2011**, *15*, 319–331. [[CrossRef](#)] [[PubMed](#)]
184. Weiss, H.M.; Cropanzano, R. Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Res. Organ. Behav.* **1996**, *18*, 1–74.
185. Kübler-Ross, E.; Kessler, D. *On Grief and Grieving: Finding the Meaning of Grief through the Five Stages of Loss*; Simon and Schuster: New York, NY, USA, 2014.
186. Bonn, S.A. Fear-Based Anger Is the Primary Motive for Violence: Anger Induced Violence Is Rooted in Fear. Psychology Today 2017. Available online: <https://www.psychologytoday.com/blog/wicked-deeds/201707/fear-based-anger-is-the-primary-motive-violence> (accessed on 25 June 2018).
187. Moss, R. Creative AI: Software Writing Software and the Broader Challenges of Computational Creativity. 2015. Available online: <https://newatlas.com/creative-ai-computational-creativity-challenges-future/36353/> (accessed on 30 August 2018).
188. Colton, S. Creativity versus the perception of creativity in computational systems. In Proceedings of the AAAI Spring Symposium: Creative Intelligent Systems, Palo Alto, CA, USA, 26–28 March 2008; p. 8.
189. Diedrich, J.; Benedek, M.; Jauk, E.; Neubauer, A.C. Are Creative Ideas Novel and Useful? *Psychol. Aesthet. Creat. Arts* **2015**, *9*, 35–40. [[CrossRef](#)]
190. Ritchie, G. Some Empirical Criteria for Attributing Creativity to a Computer Program. *Minds Mach.* **2007**, *17*, 67–99. [[CrossRef](#)]
191. Kahn, P.H., Jr.; Kanda, T.; Ishiguro, H.; Gill, B.T.; Shen, S.; Ruckert, J.H.; Gary, H.E. Human creativity can be facilitated through interacting with a social robot. In Proceedings of the Eleventh ACM/IEEE International Conference on Human Robot Interaction, Christchurch, New Zealand, 7–10 March 2016; IEEE Press: Piscataway, NJ, USA, 2016; pp. 173–180.
192. Benjamin, C.L.; Puleo, C.M.; Settipani, C.A.; Brodman, D.M.; Edmunds, J.M.; Cummings, C.M.; Kendall, P.C. History of cognitive-behavioral therapy in youth. *Child Adolesc. Psychiatr. Clin. N. Am.* **2011**, *20*, 179–189, doi10.1016/j.chc.2011.01.011. [[CrossRef](#)] [[PubMed](#)]
193. David, D.; Szentagotai, A.; Eva, K.; Macavei, B. A synopsis of rational-emotive behavior therapy (REBT); fundamental and applied research. *J. Ration. Emot. Cognit. Behav. Ther.* **2005**, *23*, 175–221. [[CrossRef](#)]
194. Young, J.E.; Klosko, J.S.; Weishaar, M.E. *Schema Therapy: A Practitioner's Guide*; Guilford Press: New York, NY, USA, 2003.
195. Walker, L.G.; Walker, M.B.; Ogston, K.; Heys, S.D.; Ah-See, A.K.; Miller, I.D.; Hutcheon, A.W.; Sarkar, T.K.; Eremin, O. Psychological, clinical and pathological effects of relaxation training and guided imagery during primary chemotherapy. *Br. J. Cancer* **1999**, *80*, 262. [[CrossRef](#)] [[PubMed](#)]
196. Wegner, D.M.; Schneider, D.J.; Carter, S.R.; White, T.L. Paradoxical effects of thought suppression. *J. Pers. Soc. Psychol.* **1987**, *53*, 5. [[CrossRef](#)] [[PubMed](#)]
197. Kaplan, F.F. Art-based assessments. In *Handbook of Art Therapy*; Guilford Press: New York, NY, USA, 2003; pp. 25–35.
198. Dan-Glauser, E.S.; Scherer, K.R. The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behav. Res. Methods* **2011**, *43*, 468. [[CrossRef](#)] [[PubMed](#)]
199. Kurdi, B.; Lozano, S.; Banaji, M.R. Introducing the open affective standardized image set (OASIS). *Behav. Res. Methods* **2017**, *49*, 457–470. [[CrossRef](#)] [[PubMed](#)]
200. Bradley, M.M.; Lang, P.J. The International Affective Picture System (IAPS) in the study of emotion and attention. In *Handbook of Emotion Elicitation and Assessment*; Coan, J.A., Allen, J.J.B., Eds.; Oxford University Press: Oxford, UK, 2007; pp. 29–46.
201. Kralj Novak, P.; Smailović, J.; Sluban, B.; Mozetič, I. Sentiment of Emojis. *PLoS ONE* **2015**, *10*, e0144296. [[CrossRef](#)] [[PubMed](#)]

