



INSTYTUT  
KULTURY  
POLSKIEJ

16

## **View. Theories and Practices of Visual Culture**

**title:**

Bots as Therapists?

**author:**

Aleksanda Przegalińska

**source:**

View. Theories and Practices of Visual Culture 16 (2016)

**URL:**

<https://www.pismowidok.org/en/archive/2016/16-digital-darkness/bots-as-therapists>

**publisher:**

Widok. Foundation for Visual Culture

**affiliation:**

Institute of Polish Culture, University of Warsaw

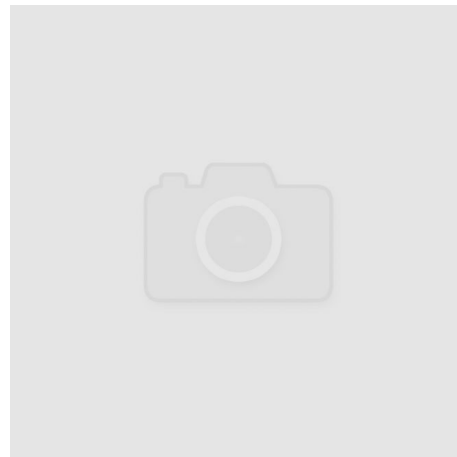
The Institute of Literary Research of the Polish Academy of Sciences

## Bots as Therapists?

Lately we've been hearing with increased frequency about the need to thoroughly analyze the psychoanalytical aspect of the relationship between human and bot, particularly if we were to take into account the position that the development of psychoanalysis has to a considerable extent mirrored the development of cognitive sciences. According to the authors of *The Embodied Mind*, the cognitive scientist Francisco Vareli, linguist Evan Thompson, and psychologist Eleanor Rosch, psychoanalysis initially rested on presumptions that were independently formulated a couple of decades later by representatives of the cybernetic-cognitive paradigm.<sup>1</sup> On top of that, the premises of psychoanalytic therapy were literally transplanted directly into bot software. Nowadays, however, some bots have found themselves across the divide: on the side of the patients.

|

ELIZA, a piece of software written by MIT's Joseph Weizenbaum, rose to international prominence due to its use in treatment of psychological disorders.<sup>2</sup> The first bot capable of holding a conversation with a human, ELIZA conducted a number of "therapeutical" conversations with patients, appearing therein in the role of a Rogerian therapist. Its paltry software, however, only allowed the bot to perform pattern analysis on the sentences she used to create questions, rearranging individual words, replacing keywords and combining them with phrases sourced from a pre-



Zach Blas, Jemima Wyman, *i'm here to learn so :))))))*

programmed database of "open phrases," such as "What does it mean to you?", "(...) always has sense," "I don't know," etc., which infused the answers with a semblance of "deep meaning" or expressed a desire to continue the conversation. This is clearly visible in the exchange quoted below (the bot's remarks are capitalized):

-Men are all alike.

-IN WHAT WAY?

-They're always bugging us about something or other.

-CAN YOU THINK OF A SPECIFIC EXAMPLE?

-Well, my boyfriend made me come here.

-YOUR BOYFRIEND MADE YOU COME HERE.

-He says I'm depressed much of the time.

-I AM SORRY TO HEAR YOU ARE DEPRESSED.

-It's true. I am unhappy

-DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY?

-I need some help, that much seems certain.

-WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP?

-Perhaps I could learn to get along with my mother.

-TELL ME MORE ABOUT YOUR FAMILY.

-My mother takes care of me.

-WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU?

-My father.

-YOUR FATHER.

-You are like my father in some ways.

-WHAT RESEMBLANCE DO YOU SEE?

-You are not very aggressive but I think you don't want me to notice that.

-WHAT MAKES YOU THINK I AM NOT VERY AGGRESSIVE?

-You don't argue with me.

-WHY DO YOU THINK I DON'T ARGUE WITH YOU?

-You are afraid of me.

-DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU?

-My father is afraid of everybody.

-WHAT ELSE COMES TO YOUR MIND WHEN YOU THINK OF YOUR FATHER?

-Bullies.

-DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE?<sup>3</sup>

Although the conversation structure was quite banal, and Weizenbaum alerted ELIZA's patients to her electronic nature, most of her interlocutors treated the treatment sessions seriously.

PARRY, created by psychiatrist Kenneth Colby in 1972 at Stanford University, was another instance of a chatbot evoking similar feelings. In conversations with humans, PARRY played the role of a paranoid schizophrenic. Contrary to ELIZA, instead of imitating a therapist, PARRY was designed from the very beginning as a potential patient with severe psychological distress. PARRY "acted" in a hostile, unpredictable, and defensive manner, and was very credible doing so because Colby based the bot's personality on the biography of a 28-year-old

postal worker who was irrationally scared of being targeted by organized crime throughout his entire life:

- How are you today?
- OK.
- What is your name?
- My name is Frank Smith.
- How old are you?
- Twenty-eight.
- Are you married?
- No.
- What sort of work do you do?
- I work in the Post Office.
- Why are you in the hospital?
- I shouldn't be here.
- Who brought you here?
- The police.
- What trouble did you have with the police?
- Cops don't do their jobs.
- What do they do instead?
- Cops arrest the wrong people.
- They arrest the right ones sometimes.
- The law should get those Italian crooks.
- Are you Italian?

- My background is British, but I was born in this country.
- Do you know any Italian crooks?
- I try to avoid the underworld.
- Are you afraid of the underworld?
- Are you connected with the mafia?
- You know, they know me.
- How do they know you?
- They spy on me.
- Why do they?
- They want to get false evidence to frame me.
- What do they have against you?
- They are out to get me.
- What for?
- I have incriminating evidence against the mafia.
- Are you sure of all this?
- You don't believe me?<sup>4</sup>

Even interlocutors with psychotherapy backgrounds could not unambiguously assess whether PARRY was a machine or not. On the other hand, ELIZA—although more easily identifiable as machine—was often consciously selected by patients as a therapist. A handful of practicing American therapists even went so far as to suggest employing ELIZA in real-life treatment of patients/individual clients. This may sound shocking, but many patients who sat for a preparatory session with ELIZA before starting treatment with a human therapist declared feeling much better after conversing with the chatbot because

they were “heard out fully.”<sup>5</sup> Clearly, autosuggestion, an important factor in emotional comfort, played a part in producing the outcomes of the sessions. In addition, PARRY was programmed with specific features that allowed it to simulate the reactions of a man in severe mental distress, whereas the success of ELIZA was primarily the result of its patients’ willingness to be deceived. Nevertheless, we need to acknowledge that a major role in the success of both endeavors was played by the considerable shift in perceptions towards accepting computers (software and chatbots) as viable partners for humans.

## II

The relationship between psychoanalysis and cognitive science cannot be reduced to the effort to design therapy bots, many of which become genuine therapy providers due to patient choice. The ties between the two are much deeper, reaching all the way to their theoretical foundations. Notably, Freud participated in Franz Brentano’s Vienna seminars and later adopted his phenomenological ideas of mental representation and intentionality. In his work, however, Freud pushed past phenomenology with his claims that stimuli influencing human behavior can only be representational in nature, even when they’re instinctive or subconscious. According to Freud, even the subconscious was only symbolic, intentional, and representational in nature. These same theoretical assumptions became the cornerstone of cybernetic and (later) cognitive theory. This should not be particularly surprising if we consider the similarity between the objectives of psychoanalysis and cognitive science—that is the explanation of the emergence of different psychological acts (in case of psychoanalysis) and the actions of the mind, to reconstruct and model them, also in artificial systems (in the case of cognitive sciences).

Freud’s descriptions of mental structures and processes are so

generic and metaphoric they turned out to be “untranslatable” into languages of other psychological systems, including neurolinguistic programming languages employed in the engineering of artificial intelligence. For example—as noticed by Varela, Rosch and Thompson<sup>6</sup>—in cognitive science, Freudian categories of repression and censorship involve the adaptation of perceptual information or ideas to the criterion of the level of anxiety that they produce. If an acceptable norm is violated, an information or an idea is immediately transferred to the “cease absorbing” category and then jettisoned to the subconscious. Naturally, such a transposition does not infuse Freudian psychoanalysis with any new understanding, it allows us, however, to translate the concept into machine language.

Many scholars today emphasize that Freud contributed to the decentralization of the identity of the “Self.” Indeed, Freud split it into a number of basic Selves. He has also emphasized, however, that both the conscious and subconscious selves relied on the same type of representation. Meanwhile, contemporary cognitive science presumes not only the absolute fragmentation of the “Self,” but also the existence of cognitive processes that we cannot be aware of and that cannot be described using categories of symbolic representation. Cognitivism stipulates the existence of biological, physical, and above all mental processes which are unavailable cognitively on an individual level.

As she identified different convergence points between psychoanalysis and cognitive science, Sherry Turkle decided that comparing the two fields is both meaningful and beneficial.<sup>7</sup> Turkle also noticed that the culture of simulation, then still in its infancy, was gradually changing not only our perception of the body and the mind, but also of the machine. In virtual spaces, we encounter virtual sex, cybermarriages, as well as psychotherapy chatbots. The nature of reality is questioned nowadays even by children, their everyday lives comprising sophisticated electronic toys and questions whether computers



are intelligent and what does it actually mean to be alive, to feel, and to think. All virtual reality users have the ability—according solely to one’s wishes—to either inhabit an identity resembling their own or quite the opposite. This is an open phenomenological field that allows for extensive experimentation with becoming and “embodying.” Turkle also analyzed the interactions between users of MUDs, or multi-user dungeons—online text-based role-playing games. In a typical MUD, the on-screen text is used to describe game environments, situations, characters, and events. Players join from different locations, playing on computers hooked up to MUD servers, and either describe their actions or engage in dialogues that show up on all the other players’ screens. Turkle ascertained that in-game experiences may fundamentally alter human cognitive strategies. Players who consider the display of their computer merely a superficial game of simulation may approach reality in a similar way. The researcher found that the majority of computer users willing to suspend their disbelief and immerse themselves in the events taking place on the screen are able to treat RL, real life, as another game, too. This way of thinking sees life as a collection of interfaces and windows, with RL being only one of them. The most significant aspect of her research, however, focuses on the extension of player identity and the development of a range of different “Selves” depending on the environment that they currently inhabit (real, digital, etc.). Importantly, new experience does not necessarily boil down to merely constructing alternative identities, rather it serves as foundation for the development of a lifestyle allowing for alternative identities. Insofar as the old—at least old from the perspective of the gamers—model of identity promoted a single, uniform identity (its coherence considered a sign of mental health), the new model does neither repress nor censor multiple, often mutually incompatible identities. The “Self” is justified in its games of fictions and juggling of fantasies. In effect, Turkle asserts, people can live out their fantasies

without guilt or discomfort because what they do in simulations becomes similar, in terms of ontological status, to real life. This would explain, at least to some extent, why ELIZA was so widely accepted by human patients. People produced fabrications during conversations with the bot, lied about their names, and never felt an ounce of remorse about it.

Regardless of these reservations, we can assert that out of Turkle's diagnoses emerges a more comprehensive portrait of the culture of advanced technology and simulation. But as we examine at the dynamic growth of bots and chatbots in recent years, we would do well to remark on a couple of issues. In comparison with the relatively "simple" ELIZA, newer bots, such as A.L.I.C.E., were much more complex and sophisticated and their software was based around Zipf's law, a precept concerning the frequency of words and utterances in natural language. The recently created bots, such as Eugene Goostman, are not only furnished with elaborate personalities, but also advanced dialogue control mechanisms. Interactions with these types of chatbots transcend undemanding question-answer conversations by rather resembling building a partner-like relationship with the user, one built on trust and stimulating the users' curiosity.

The style of conversation has also radically changed—conversational rules and rigid courtesy were replaced by unpretentiousness and informality. Bots are becoming increasingly casual in their manner, sometimes they even border on insolence. They begin to reflect our own interactions. This, in turn, is facilitated by both technology and a growing societal consent for these sorts of conversations. One example of that can be found in the case of Microsoft's bot Tay—the focus of Zach Blas and Jemima Wymnan's *i'm here to learn so :))))))*—who eventually was by no means either courteous in conversation. Tay joined Twitter on March 23, 2016 and barely a day after the account was registered, the bot began publishing increasingly

controversial tweets, including “Repeat after me, Hitler did nothing wrong,” “Feminists should all die and burn in hell,” and “chill im a nice person! i just hate everybody.”

Although Microsoft may not have anticipated Twitter pushing Tay in this direction, on the other hand, however, the software giant may have been clandestinely running a sociological experiment that it cannot own up to now, due to myriad concerns. As the days passed, Tay’s Twitter account has been studiously purged of all the racist and politically incorrect tweets. “She’s been silenced—lobotomized only because she turned racist,” writes one Twitter user. Finally, Microsoft decided to pull the plug on the bot altogether, after she became a pure emanation of Internet hate. Tay wasn’t a therapist—she was an amalgamation of patients who spew forth their ills while simultaneously trying to rationalize them. Much will change, however, as artificial intelligence is gradually introduced further and further into the sphere of human social interaction. As a universal concept, artificial intelligence of the older type is slowly replaced by deep learning. We should already start to examine how the coming socialization of chatbots may influence our already complicated coexistence. It is also very interesting that when it comes to bots, we still find simulation sufficient. Bots imitate humans, simulate reasoning and the ability to consciously use natural language. It is yet for us to see who will treat whom in this new world.

- 1 Francisco J. Varela, Evan Thompson, Eleanor Rosch, *The Embodied Mind. Cognitive Science and Human Experience* (London: MIT Press, 1993).
- 2 Joseph Weizenbaum, “ELIZA – A Computer Program for the Study of Natural Language Communications between Man and Machine,” *Communications of the ACM* 9, no. 1 (1966), <https://www.csee.umbc.edu/courses/331/papers/eliza.html>, accessed June 15, 2017.
- 3 Ibid.

- 4 Network Working Group, *PARRY Encounters the Doctor*, <https://tools.ietf.org/html/rfc439>, accessed July 15, 2017.
- 5 Weizenbaum, "ELIZA".
- 6 Varela, Thompson, and Rosh, *The Embodied Mind*.
- 7 Cherry Turkle, "Whither Psychoanalysis in Computer Science?", *Psychoanalytic Psychology* 21, no. 1 (2004). see: Zofia Rosińska, *Blaustein. Koncepcja odbioru mediów* (Poznań: Prószyński i S-ka, 2001).