The "Eliza Effect" And Its Applications In Psychotherapy

Ioana A. Cristea
University of Pisa, Italy & Babes-Bolyai University,Romania
Madalina Sucala
Babes-Bolyai University, Romania
Daniel David
Babes-Bolyai University, Romania & Mount Sinai School of Medicine, USA

The Eliza Effect

- "The susceptibility of people to read far more understanding than is warranted into strings of symbols — especially words — strung together by computers" (Hofstadter, 1996)
- Notable for occurring even when users of the system are aware of the determinate nature of output produced by the system.
- From a psychological standpoint, the ELIZA effect is the result of a subtle cognitive dissonance between the user's awareness of programming limitations and their behavior towards the output of the program (Ekbia, 2008).

Alan Turing and the Turing test

- In 1950, in the article *Computing Machinery and Intelligence*, which appeared in the philosophy journal *Mind*, Alan Turing asked the question "Can a Machine Think?"
- He answered in the affirmative, but a central question was: "If a computer could think, how could we tell?"
- Turing's suggestion was, that if the responses from the computer were indistinguishable from that of a human, the computer could be said to be thinking.

Chatterbots (Chatbots)

- Computer programs designed to simulate an intelligent conversations with one or more human users via auditory or textual methods.
- Recently, chatbot-like methods have been used for practical purposes such as online help, personalized service, or information acquisition, the programs functioning like conversational agents.
- Simplicity of the algorithms used. Although many chatbots do appear to interpret human input intelligently when generating their responses, many simply scan for keywords within the input and pull a reply with the most matching keywords, or the most similar wording pattern, from a textual database.

ELIZA

- ELIZA is a computer program by Joseph Weizenbaum, which, designed in 1966, generally recognized as the first chatbot.
- Eliza parodies a Rogerian therapist, largely by rephrasing many of the patient's statements as questions and posing them to the patient. Thus, for example, the response to "My head hurts" might be "Why do you say your head hurts?". The response to "My mother hates me" might be "Who else in your family hates you?"
- In 1976, Weizenbaum wrote: "*I was startled to see how quickly and how very deeply people conversing with Eliza became emotionally involved with the computer and how unequivocally they anthropomorphized it.*"

ELIZA

- ELIZA was named after Eliza Doolittle, a working-class character in George Bernard Shaw's play Pygmalion, who is taught to speak with an upper class accent.
- http://www.abcguides.com/abcsoftware/eliza_js_noframe.htm
- Sample interaction with ELIZA

Human: Well, my boyfriend made me come here.
ELIZA: Your boyfriend made you come here?
Human: He says I'm depressed much of the time.
ELIZA: I am sorry to hear you are depressed.
Human: It's true. I'm unhappy.
ELIZA: Do you think coming here will help you not to be unhappy?

The Loebner Prize for Artifical Intelligence

- http://loebner.net/Prizef/loebner-prize.html

- In 1990 Hugh Loebner agreed with The Cambridge Center for Behavioral Studies to underwrite a contest designed to implement the Turing Test. Dr. Loebner pledged a Grand Prize of $100,000 and a Gold Medal for the first computer whose responses were indistinguishable from a human's.

- Each year an annual prize of $2000 and a bronze medal is awarded to the **most** human-like computer. The winner of the annual contest is the best entry relative to other entries that year, irrespective of how good it is in an absolute sense.

- Loebner prize Gold Medal

"Eliza like" developments

- A.L.I.C.E. (Artificial Linguistic Internet Computer Entity)
- http://www.pandorabots.com/pandora/talk?botid=f5d922d97e345aa1 (online since 1995; Loebner prize winner 2000, 2001, 2004)
- Based on a complex set of rules ("if-statements") that govern its response to a question

"Eliza like" developments
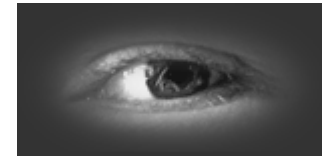
- JABBERWACKY
- http://www.jabberwacky.com/ (online since 1995, Loebner prize winner 2005, 2008)
- Stores everything everyone has ever said, and finds the most appropriate thing to say using contextual pattern matching techniques. In speaking to you it uses just that learnt material, and borrows a little bit of your intelligence as it learns more.
- With no hard-coded rules, it relies entirely on the principles of feedback.
- "Conversational wikipedia"

And finally...

- Talk to God
  - http://www.alicebot.org/igod/

Implications for Psychotherapy

- With the growing role of the computer-assisted psychotherapy, the field needs to clarify which are the most relevant therapeutic factors to be included in computer assisted therapies.
- In present, embodied computer agents are designed with more and more human characteristics, like voice, appearance and even the capacity for emotional recognition and expression (Brave, Nass, & Hutchinson., 2005).
- Systems have been built that are able to provide emotional support for their users (Klein, Moon, & Picard, 2002).

Implications for Psychotherapy

- These developments are bound to be linked with the practice of psychotherapy especially since they help overcome two of the major problems people have when considering therapy: costs and the loss of anonymity.
- Since computer-based systems will start playing an increasing role in the practice of psychotherapy, it is relevant to have studies that compare the performance of these systems with trained human therapists on a number of dimensions, which are important for the effectiveness of psychotherapy (common and specific factors).

Implications for Psychotherapy

- The issue is particularly germane both from the point of view of the user (patient), but also from the one of other therapists, who could assess the quality of the interaction on therapeutically relevant dimensions.
- It is also an important aspect for the robotics or AI industry, which are getting more involved in producing intelligent agents, for a number of purposes, including for psychotherapy.

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- Our study is an exploratory one, designed to compare the performance of the ELIZA program with trained human therapists on a number of dimensions, important for the effectiveness of psychotherapy.
- We aimed to investigate how therapists evaluate and compare the interaction between a patient and a real therapist and the interaction between the same patient and a simple software program that simulates therapeutic interventions (ELIZA).

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- The therapist used specific cognitive-behavioral techniques (e.g. problem solving, restructuring), in addition to the techniques used for developing and maintaining the therapeutic relation (e.g. empathic reflection, summarization), thus employing both common and specific factors in the interaction.

- ELIZA, which simulates the interaction with a real therapist, is based on simple pattern recognition and substitution of key words into standardized questions or comments (see Weizenbaum, 1976). Therefore it made use only of common factors in the intervention.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- To our knowledge, this is the first study to experimentally compare a computer-based therapeutic agent (the Eliza program) with a trained human therapist, from an expert point of view (not simply regarding the patient's subjective perception of the interaction).

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- **Procedure**
  - A real client of one of the authors of the study underwent two short interventions on the same problem (chosen by her).
  - The client was told that this was a study involving two therapists, one of whom was her usual therapist.The other one was in a different location and could be contacted by the means of a chat system.
  - After 15 minutes the interaction was stopped and the client entered the session with her therapist. The same problem was discussed for 15 minutes at the beginning of the session.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- **Procedure**
- A rating scale to evaluate therapist performance was constructed.
- The items referred to common factors variables, which have been shown to be important in constructing an effective therapeutic relationship (Lambert & Barley, 2002): (1) empathy; (2) unconditional acceptance of client/patient; (3) collaboration; (4) attention to the client/patient; (5) care about what the client says, as well as to variables having to do with specific factors, regarding (1) the efficiency of the discussion, (2) the evaluation of the therapist's approach, and (3) overall performance of the therapist.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- **Procedure**
- The transcripts together with the scale were sent to Romanian and foreign therapists.
- The scale contained the instruction that the attached materials were two interactions between the same patient and two different therapists.
- The respondents were also asked to mention the form of therapy practiced and their level of expertise (expert - Ph.D or attained supervisor status - or novice). The instruction was given that only clinical psychologists or therapists working with patients, in or out of supervision, should take part in the study.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- ## **Participants**

A number of 138 therapists responded to our query.

- Nationality: 116 (84.1%) were Romanian and 22 (15.9%) foreign.
- Gender distribution: 28 males (20.3%), 97 female (70.3%), and 13 (9.4%) did not report their gender.
- Level of expertise: 87 participants (63%) listed themselves as novice therapist, while 36 participants (26.1%) as experts and 15 participants (10.9%) did not report their level of expertise.
- Form of therapy practiced: 59 participants (42.8%) reported practicing cognitive-behavioral therapy (CBT), 59 participants (42.8%) other forms of therapy and 20 participants (14.5%) did not declare the type of therapy they practiced.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- ## **Results**

- Significant differences between the Eliza program and the real therapist on each of the aspects (*p<.01)

Variable	Eliza Mean (SD)	Human Mean (SD)	t value
Efficiency of discussion	1.54 (0.94)	3.87 (0.85)	-20.264*
Unconditional acceptance	1.92 (1.03)	4.30 (0.78)	-19.729*
Collaboration	2 (1.02)	4.30 (0.73)	-18.894*
Empathy	1.83 (1.00)	4.13 (0.85)	-18.333*
Attention to the client	2.13 (1.14)	4.39 (0.86)	-17.254*
Care about what the client says	2.10 (1.10)	4.23 (0.82)	-16.421*
Correct approach of the problem	1.60 (0.88)	3.96 (0.90)	-19.435*
Overall performance	1.50 (0.88)	4.10 (0.85)	-22.208*

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- ## **Results**
- We also wanted to see if there were some variables that could act as moderators for the proficiency to distinguish the Eliza program from the real therapist.
- The sex of the therapists (male versus female), their nationality (Romanian versus foreign), as well as their self-declared level of expertise (expert versus novice) did not have any influence on their proficiency in distinguishing the Eliza therapist from the human CBT therapist (all ps >.05).

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- **Results**
- However, the type of therapy practiced did have an influence on some of the dimensions taken into account.
- More specifically, there were significant differences between CBT experts and the others, favoring CBT experts, on the dimensions regarding the efficiency of the discussion ($t(86.18)= 2.89$, $p<.05$), the degree of collaboration exhibited ($t(102.5)= 2.69$, $p<.05$), the correct approach of the problem ($t(87.5)= 4.73$, $p<.05$) and the overall appreciation of the therapeutic performance ($t(90.4)= 3.41$, $p<.05$).

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- **Discussions and implications**
- The subjects who evaluated the transcripts from the two therapists (Eliza software and human therapist) seemed to view them both as human therapists. At no point did any of them mention that the interaction with therapist 1 (the Eliza program) displayed unusual features or that it was a strange, unnatural therapist-patient interaction.
- The perceived difference between the Eliza program and the human therapist seems to lie in the quality of the performance, and not in some intrinsic features of either.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- **Discussions and implications**
- We envisioned an interesting question might be whether there are any therapist related variables that can moderate the ability to distinguish between the computer agent and the human therapist, on some of the dimensions we took into account.
- The variables we took into account were selected based on therapist outcome studies, which tried to identify therapist-related values that may affect therapy outcomes (e.g. Orlinsky & Howard, 1980; Hersoug, Høglend, Monsen, & Havik, 1991).

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

- ## **Discussions and implications**
- We found that the sex of the therapists, their nationality, or their self-declared level of expertise had no impact on their capacity to distinguish the Eliza program from the human on any of the variables taken into account.
- What seemed to have a selective impact was the form of therapy they declared to practice. This is consistent with the idea that, while the aspects regarding the construction of a good therapeutic relationship are common across different forms of psychotherapy, things look different when it comes to efficiency, approach of the problem, and other variables which have to do with specific factors in psychotherapy.

Cristea, I., Sucala, M., David, D. (submitted). The "Eliza Effect": The Anthropomorphic Tendency In Computer Aided Psychotherapy

Final considerations

- The research on computer based therapists is scarce. We have not been able to come across studies comparing them with human therapist, from what is considered a professional/expert standpoint.
- However Eliza is not among the complex therapeutic computer based agents. Its principles are pretty basic and it operates following simple script rules.
- A useful next step would be to compare more complex computer-based therapy agents with human therapists in order to see if this pattern of results remains consistent.

Thank you for your attention