

# Cognitive Psychology: A Solution to the Black Box Problem in Artificial Intelligence (AI) - Analysis with Respect to the Tamil Film ‘Enthiran’

OPEN ACCESS

Volume: 12

Special Issue: 1

Month: April

Year: 2025

E-ISSN: 2582-0397

P-ISSN: 2321-788X

Citation:

Shanmugaa Bharathi, SR.  
“Cognitive Psychology:  
A Solution to the Black  
Box Problem in Artificial  
Intelligence (AI) - Analysis  
with Respect to the Tamil  
Film ‘Enthiran.’” *Shanlax  
International Journal  
of Arts, Science and  
Humanities*, vol. 12,  
no. S1, 2025, pp. 115–19

DOI:

[https://doi.org/10.34293/  
sijash.v12iS1-Apr.8955](https://doi.org/10.34293/sijash.v12iS1-Apr.8955)

**SR. Shanmugaa Bharathi**

23ENG431, II M.A. English Literature  
Department of English and Foreign Languages  
Bharathiar University, Coimbatore

## Abstract

Psychology is generally called the study of the human mind and has many aspects to it including the developments we undertake in technology with the aid of the subject. Artificial intelligence is the representation of the human mind and we have created artificially intelligent robots as an incarnation of the human mind that modern humans do not take the effort to use their minds for most of the work. Rightly so, they are called the counterpart of humans. Though artificially intelligent beings lack emotional intelligence that humans have, they are made with the maximum possibility to mutate and adapt themselves into human society. Artificial intelligence is now capable of taking decisions and portraying emotions that have consequences to human beings due to the significance of their relationship in today's world. To arrive at such decision making, some areas of cognitive psychology like learning and memory, social cognition, complex attention and problem-solving abilities can be used. The use of such areas of cognitive psychology can be applied to Director Shankar's famous Tamil film 'Enthiran'. The film portrays the introduction of a robot developed by a scientist and its integration with sapiens as it develops its personality gradually. This paper attempts to study the use of the above-mentioned cognitive areas in pursuing a solution to the black box problem in artificial intelligence.

**Keywords:** Artificial Intelligence, Cognitive Psychology, Black Box Problem, Enthiran.

## Introduction

Artificial intelligence is a representation of the human mind. Using artificial intelligence, we have created robots as an incarnation of the human mind that we modern humans are constantly striving to work and think lesser while allowing and expecting the robots to do all the work. Rightly called the 'human counterpart', they are also fed with cognitive abilities of creativity, learning, thinking and problem-solving capacities (Kaur). Artificial intelligence is developed to such great heights that they autonomously make decisions and alter their responses according to the feedback of the respective decision's interpretation. Emotions are also programmed into the AI database to provide customised responses. But the processing of the emotions as a response to a given situation is unknown as the black box is

designed in such a fool-proof way to prevent the process from being visible and interpreted. This is termed as the 'Black box problem' in AI.

In order to clarify the source of such emotions and solve the problem of autonomous decision making, there is a need to analyse the functioning of artificial intelligence using the theories of cognitive psychology that are used to analyse the human mind. Cognitive psychologists believe that there is an internal process called 'mediational process' that happens internally before any response to a stimuli is received. Cognitive psychology advocates the understanding of the mediational processes to understand behaviour and emotions of a person (Jordan). The same is being applied to understand the expression of emotions by artificial intelligence. This paper focuses on the use of cognitive psychology to analyse the emotional quotient and decision-making capabilities of robots. This will be done by analysing the robot featured in the movie Enthiran.

The film by Shankar is about a male robot called Chitti and his survival and role in human society. The various features that Chitti can perform are glorified throughout the film, which includes its decision-making abilities, creative thinking, emotional balance, and personality development. Chitti becomes the hero and the villain due to his ability to think autonomously and to pick up emotions quickly from its database. He is the hero in the beginning, but as the film progresses, it can be seen that he gains the ability to make his own decisions and alter responses to any question asked although not included in the programming. In the process of doing so, he makes disastrous decisions to the extent of causing harm to his very own creator and caretakers. The humans appear to be clueless about Chitti's strong human-like feelings and as the story progresses, Chitti is discarded as his abilities become dangerous to the very humans who created him, albeit temporarily. It can then be seen that Chitti is recovered and his abilities are used negatively by the antagonist. The film ends with Chitti being forced to be dismantled due to the fear of the black box problem. This paper therefore attempts to prove the use of cognitive psychology to arrive at solutions for the problem.

### Literature Review

C.F.Schmidt, N.S.Sridharan&J.L.Goodson(1978), inferred from their study on artificial intelligence (AI) that AI can be applied on psychology by looking at the knowledge sources of a plan recognition process that was used to define the application of artificial intelligence on psychology. The experiments were done on humans and then applied on AI. The theory dealt with sequences that account for the goals or responses in humans and likewise in AI.

According to Estes, cognitive psychology investigates every aspect of information processing, acquisition of concepts about the world, their utilisation in decision-making and carrying out activities. It also investigates any processing, not necessarily only of the human mind. This study has used this theory of the use of cognitive psychology in identification of decision making and emotional expression by artificial intelligence as opposed to the human mind.

J.Ericet al (2020), explained the complexity of AI in today's world and published a review of the data programme used in the black box along with the results of experiments they had conducted on AI using basic psychology. They also invite more cognitive psychologists to conduct additional experiments in the field to provide more explanation to the black box problem.

### Discussion

#### Digital Anthropology and Artificial Intelligence

Digital anthropology is the process of analysing relations between humans and digital technology (Miller). Technology works on artificial intelligence with the black box and humans have seemed to readily welcome any change that enhances technology. Artificial intelligence products like Alexa and automatic floor cleaners that have a black box designed for smoother human interaction

have become an inevitable part of human life today and so have robots (Kaur). It is mainly because they are able to reciprocate the emotions that humans either need or want to see. Humans have also started accepting robots in their lives as seen in the film *Enthiran* that is taken for study.

While using the fields of digital anthropology in the film, it can be seen clearly that in the initial parts of the film, everyone who knows Chitti, seem to have a humane relation with him. Even though Chitti is programmed through Machine Learning, it can be seen that the humans assign it the male gender, family relations, for instance, Amma. It is able to make decisions, relate emotionally and provide answers to any question. It is to be noted here, that this was done to serve humans, but towards the end, Chittius uses the programming in ways that makes tables turn. This is proof that cognitive ability analysis, personality traits, creative thinking and so on can be applied to AI as well.

### **Black Box Problem in Artificial Intelligence:**

Black box in AI is "...where AI produces insights based on a data set, but the end-user doesn't know how." It is basically an impenetrable system in AI used for the decision making process, but is unknown to all (Reed). It decides whether a particular symptom is the disease, whether a student is capable of being an advocate in their future, and many more important things. In these circumstances, it becomes an issue of how the AI knows if a thing is true or not or whether an answer is right at a particular situation (von Eschenbach). AI functions mysteriously in making such decisions and formulating emotions and the process is unknown even to creators of the black box (Rudin&Radin). Similarly, in the movie *Enthiran*, Chitti's emotions surprise the audience with its emotional intelligence.

To provide a solution to this problem, programmers have come up with the concept of "Explainable AI", in which the results provided by the AI are understood and explainable by humans (Taylor and Taylor). But this feature is not fully compatible for commercial use as of today. In such circumstances, cognitive psychology can be used to try to understand how these decisions and emotions are processed by AI. As cognitive psychology deals with understanding the internal mediational process of the human mind that lead to certain behaviour, it can be applied to AI. There's more ground to support such application as AI is the representation of the human mind, designed initially as the human counterpart. Therefore, it can be applied on Chitti in the film *Enthiran* as well.

### **Findings**

#### **Cognitive Psychology as a Problem Solver**

Cognitive psychology researchers Pooja K. Agarwal & Henry L. Roediger III used results from their cognitive psychology studies and developed better practices to encourage more efficient learning in the classroom. They looked into how memory and grasping power work in order to enable students to produce results and solutions to the questions accordingly. Similarly, the memory and retention of the black box can be analysed by using the same experiments used on humans to identify what determines the resultant decisions made and emotions shown. Behaviour, responses and opinions that the AI produces can be considered in arriving at how the stimulus is being processed into responses in the black box (Taylor and Taylor).

This can be analysed in the film *Enthiran* by looking at Chitti's behaviour in several scenes in the film. Chitti the robot slowly develops his own answers like the answer he gives in the meeting when asked to kill his creator. He was not programmed to have understood the importance of the order and act accordingly. Nevertheless, he understands and acts accordingly, leaving the audience shook. This shows that his emotions give access to what kind of an action is taking place in its

algorithm, just like human emotion gives answers to the thoughts going on in one's mind. There are scenes that show Chitti having developed romantic affection for its creator's fiancé crossing the thin line differentiating the living from the non-living. Using the theory of cognitive psychology, his mediation process such as what he interprets of relationships, his understanding of life with someone, etc., show to an extent the cause for Chitti developing such intimate feelings for Sana the fiancé. Hence, cognitive psychology can be accepted without doubt as a problem solver in this case.

### Relevance and Inference

The film portrays Chitti as an independent robot and shows implicitly the consequences of absence of an explanation to its emotions and decisions. If Chitti was given a memory experiment, his ways of interpreting and retaining what has been received could be found out. For example, if an experiment on his memory is to be conducted to reveal how he identifies as a man, we would have to analyse how he first interprets gender, what he has understood when he was assigned it and how he has come to accept his physical appearance.

The same way, if an experiment of creative thinking is done on Chitti, the reason for his answer containing emotion and his desire to live even after his creator had abandoned him can be solved. For instance, if the experiment seems to bring out what he feels about emotion through what his data has stored as emotion, one can see how he identifies different emotions, and how he is subjected to feel and how he is replicating what he feels.

Cognitive psychologists in the process of conducting these experiments work towards the result that would help the humans better in their thinking and decision making or memory capabilities. Likewise, experiments conducted on AI can produce results that could help in better framing the black box in a way that is easier to understand than the explainable AI that is being worked on as a solution currently. Cognitive psychologists have disproved many myths and validated some of them with their results. AI emotions can also be determined using the above said methods and can be modulated to better understand humans in the upcoming future (Taylor and Taylor).

### Conclusion

Humans have contributed vastly towards the development of artificial Intelligence to perform the work that is needed, but is undesired by humans. Through this process, humans have made artificial intelligence very deep that the process they use to make decisions and formulate emotions is unknown (Rudin & Radin). This has led to a recent fear of AI replacing human labour (von Eschenbach). This process shows their development in parallel to the human mind in making decisions. Though this process is seen as an area of vast development in achieving the goals related to origins of AI, it becomes an issue of concern in places where the decisions made by the robot is in any way a factor of consequence to humans (Kaur). The absence of an explanation to the same reduces the trust on AI.

To solve this issue, the theory of mind mapping used under cognitive psychology can be used as it is an attempt to learn more about brain cells to understand their connection with the health and behaviour of persons. The data set used to program Chitti can be similarly learned to understand the connection with his behaviour and emotion. The theory of short-term memory can also be used on Chitti to determine how his memory serves as a determiner in making decisions as the results are based on what has been previously fed to the drive that contains all the data. In this way, how the memory is stored and re-used to give a response according to the situation can be analysed by looking at how long the memory is saved and its similarity with the human memory. Thus, it can be concluded that experiments of cognitive psychology used on Chitti in Enthiran helps in understanding the processes behind AI's emotional quotient and decision-making abilities.

## References

1. Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951715622512>
2. Cherry, Kendra. “Cognitive Psychology.” *Verywell Mind*, Verywell Mind, 14 Nov. 2022, <https://www.verywellmind.com/cognitive-psychology-4157181>.
3. Collins, Allan, et al. “Computing Machinery and Intelligence.” *Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence*, Morgan Kaufmann, San Mateo, California, 1989, pp. 6–18.
4. Jordan, Candina. “Cognitive Psychology: How Scientists Study the Mind.” *Web MD*, WebMD, 2022, <https://www.webmd.com/mental-health/what-is-cognitive-psychology#:~:text=Cognitive%20psychology%20is%20the%20branch,think%20the%20way%20you%20do>.
5. Kaur, Dr. Suninder. “Evolution in Human-Computer Interaction.” *UIAF, United Indian Anthropology Forum*, 16 Nov. 2021, <https://www.anthropologyindiaforum.org/post/evolution-in-human-computer-interaction>.
6. Miller, Daniel. “Digital Anthropology - Anthroencyclopedia.com.” *Digital Anthropology - Cambridge Encyclopedia of Anthropology*, University College London, 2018, <https://anthroencyclopedia.com/printpdf/462>.
7. Reed, Niamh. “The AI Black Box Problem.” *ThinkAutomation*, Parker Software, 11 Nov. 2021, [https://www.thinkautomation.com/bots-and-ai/the-ai-black-box-problem/#:~:text=The%20AI%20black%20box%2C%20then,it%20produces%20\(the%20output\)](https://www.thinkautomation.com/bots-and-ai/the-ai-black-box-problem/#:~:text=The%20AI%20black%20box%2C%20then,it%20produces%20(the%20output)).
8. Rudin, Cynthia, and Joanna Radin. “Why Are We Using Black Box Models in AI When We Don’t Need to? A Lesson from an Explainable AI Competition.” *Harvard Data Science Review*, HDSR, 22 Nov. 2019, <https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8>.
9. Schmidt, C.F., et al. “The Plan Recognition Problem: An Intersection of Psychology and Artificial Intelligence.” *Artificial Intelligence*, Elsevier, 16 June 2003, <https://www.sciencedirect.com/science/article/abs/pii/0004370278900127>.
10. Taylor, J. Eric T., and Graham W. Taylor. “Artificial Cognition: How Experimental Psychology Can Help Generate Explainable Artificial Intelligence - *Psychonomic Bulletin & Review*.” Springer, Springer US, 6 Nov. 2020, <https://link.springer.com/article/10.3758/s13423-020-01825-5>.
11. von Eschenbach, W.J. Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philos. Technol.* 34, 1607–1622 (2021). <https://doi.org/10.1007/s13347-021-00477-0>