Bartneck, C. (2023). The Dorian Gray Refutation. Behavioral and Brain Sciences, 46(e22). https://doi.org/10.1017/S0140525X22001662

The Dorian Gray Refutation

Christoph Bartneck

April 5, 2023

Abstract

Theories are an integral part of the scientific endeavour. The target article proposes interesting ideas for a theory on human–robot interaction but lacks specificity that would enable us to properly test this theory. No empirical data are yet available to determine its predictive power.

1 Disclaimer

This is an open peer commentary for the article "Social robots as depictions of social agents" in the Behavioral and Brain Sciences journal [Clark and Fischer, 2023].

2 Introduction

I wish Clark and Fischer (C&F) had given their theory a name because that would make it easier for us to talk about it. For the time being, let's call their theory the Clark and Fischer Conjecture (CFC). I then have to wonder how the CFC is better than other theories, such as the media equation. To better understand this question, we have to consider the role that theories play.

Constructing theories is an essential process in the scientific endeavour. Theories help to explain the past and predict the future. There are several criteria available to judge the value of a theory. Arguably, the most important criterion is its explanatory power.

The more observations a theory is able to accurately model, the higher its value. A theory, such as gravity, that applies to all things is considered more powerful than a theory that only applies to robots. A theory that applies to all robots is more valuable than a theory that only applies to social robots and so forth. Achieving a higher generalisability often requires the use of more abstract terms in the theory. This does, at times, lead to situations where researchers engage in discussion about semantics rather than about models of reality. The relationships of these terms are then ideally expressed using maths.

Another success criterion for a theory is the accuracy and reliability of its explanations and predictions. Lastly, the simplicity of the theory itself makes it preferable over others. Occam's Razor dictates that a theory that uses fewer concepts to model reality is preferable.

Before we start a discussion of the merits of the CFC we have to acknowledge that Nass and Reeves [1996] would probably not agree to C&F's representation of the media equation as "media = real life." This is an over-simplification. We all know that movies are just movies. While they still have the power to make us cry, we know that they are just a representation. We have little trouble experiencing this cognitive dissonance. Hence media is not exactly the same as real life. C&F seem to have employed the rhetorical straw-man technique to highlight the need for a better theory. This is unnecessary, as there is little doubt that human–robot interaction (HRI) is a new form of media that requires further attention. HRI was only in its infancy when the media equation was proposed.

Is the CFC better than other theories, such as the media equation? At this stage we cannot say. There are no experiments available yet that have demonstrated that the CFC succeeds over other theories based on the success criteria mentioned above.

This leads us to one of the main challenges of this paper. I struggle with fully understanding the CFC because of its complexity. It considers a large number of concepts such as the varieties of depictions, perspectives on depictions, character types, imagination in depictions, frames of reference, layers of activity, authorities, and emotions. Many of these concepts are then further subdivided. Varieties of depictions, for example, is subdivided into static versus dynamic depictions, staged versus

interactive depictions, and actor versus prop depictions. This part of the CFC alone could be considered a theory worthwhile of testing. With these large number of concepts it seems daring to come up with a mathematical expression for the complete CFC.

Unless we have such a mathematical expression of the CFC it will be difficult, if not impossible, to construct empirical studies to test the CFC. It could even be argued that an experiment that tests all the aspects of the CFC would be so complex that it would become highly impractical to test.

How then can we ever know that the CFC is true? How can we ever know that the CFC has more explanatory power than the theories, such as the media equation? How can we know that the predictions of the CFC's are more accurate and reliable? We may never know conclusively. What we do know is that the CFC is complex. Far more complex than "media = real life."

This complexity is not only because of the number of concepts involved, but also because of unspecific relationships between them. The most we can learn from this paper is that concept A somehow relates to concept B. What we miss are more precise predictions, such as $A = 2 \times B$. This level of specificity is necessary to fully understand the concepts and their relationships.

This does not mean that the CFC has no merit. Formulating theories is important. Most studies in HRI dissect concepts into ever finer slices of reality that they then study in isolation. Far less effort is made towards bringing all these concepts back into an overarching theory. C&F should be applauded for their effort, even if their conjecture will be revised and extended in the future.

C&F remind us of concepts that are likely to play an important role in HRI. We should consider the robot itself, its representation, and the interaction context. What exact influence they have, however, remains unclear. It would have been desirable if C&F would have given us more clues on how we can test their conjecture.

I would like to end with a quote from Oscar Wilde who fails to lose relevance even after all these years. In his novel, Dorian Gray states that, " ... no theory of life seemed to him to be of any importance compared with life itself. He felt keenly conscious of how barren all intellectual speculation is when separated from action and experiment." Let's bring CFC closer to action and experiment by constructing systematic studies that will shed light on its concepts and their relationships.

References

Herbert H. Clark and Kerstin Fischer. Social robots as depictions of social agents. *Behavioral and Brain Sciences*, 46:e21, 2023. doi: 10.1017/S0140525X22000668.

C. Nass and B. Reeves. *The Media equation*. SLI Publications, Cambridge University Press, Cambridge, 1996. RefMgr field[1]: Book, Whole RefMgr field[8]: Not in File.