The Color of Your Eyes

by Luc Steels

LOOK INTO THE BOX was an interactive installation devised by Olafur Eliasson and myself for the Musee d'Art Moderne de la Ville de Paris. It was shown there in the spring of 2002 as a part of Eliasson's solo exhibition and shown again at the Spoleto Scienza Festival in the summer of 2003.

When a viewer enters the darkened installation space, one sees in a corner a small box standing on a pole, looking a bit like a camera obscura with a hole in the middle. There is a magnifying glass before the hole and a light just on top of it. Below the box, there is a pedestal on which one can stand to look in the box. When the viewer does this, there is a clicking sound, similar to that of an old-fashioned camera taking a picture. Immediately after, the viewer's eye appears as a large projection on the opposite wall. Seeing one's eye this way is a revealing, almost shocking, experience, for there are few things more intimate and beautiful than the human eye. Soon after the eye appears, spoken sounds emanate, almost as if angels were whispering words in a strange language. One angel might say: "Ko. Lecefi. Guso," and the other one repeats it in a questioning tone: "Ko? Lecefi? Guso?" In another dialog, you might hear: "Ko. Lecefi. Lulke," or "Ko. Low. Gi."

Thanks to Eliasson's careful design, the aesthetic experience of *Look into the box* is immediate. But to gain a more profound understanding of the piece, we must observe – with a discerning eye – both inside and beyond the box. When we do so, a web of intriguing connotations and insights reveal themselves. The box contains a camera and a computer, which processes the incoming images to detect whether a human eye is

Look into the box, 2002
Dimensions variable

This photograph depicts the author in the installation standing in front of an image of his own eye captured by the camera and projected on the wall behind him. Installation view at Chaque matin je me sens différent, chaque soir je me sens le même, Musée d'Art Moderne de la Ville de Paris

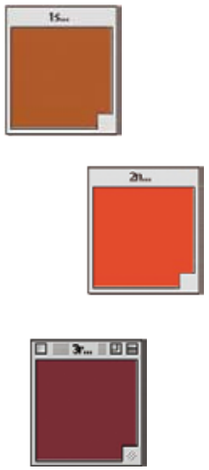
Photograph: Marleen Wynants

present. This in itself is no small feat, and it required creating adaptive-pattern-recognition algorithms that were first trained with images of human eyes in the specific context of the exhibition space.

As soon as an eye is detected, a picture is made, and this picture is "shown" to two "agents," who then play a language game describing to each other the colors of the eye they just saw. The rather astonishing point of the experiment is that the agents begin without any names and even without color distinctions. They must invent their own. Thus, "ko" or "lecefi" are examples of the kinds of color words the agents invented and used.

Agent is a term used in contemporary artificial-intelligence research to denote artificial entities that have autonomy and intelligence. They are ubiquitous software programs that may use cameras, microphones, speech synthesizers, and robot bodies to interact with the world. Many experiments are currently being done that involve populations of interacting agents, and that is the case here.

fig. 1
Above: Sample eye, pixelated eye
Below: Dominating colors of the eye



In fact, *Look into the box* was erected as both an artwork and a scientific experiment. It explores issues in the new field of semiotic dynamics, semiotics being a branch of philosophy that studies the relationships among the world, language, and the categorization of the world as expressed in language. Semiotic dynamics studies how these relationships emerge and change in a population under the influence of an evolving world, the changing constellation of the population, or simply the internal dynamics generated by grounded, situated communications. Whereas classical semiotics is mostly descriptive and intended to help the interpretation of texts or pictures, semiotic dynamics uses all the tools of modern science – including computer simulations, robotic experiments, and the mathematical techniques from complex systems science – to build a scientific theory on how language and meaning are possible.

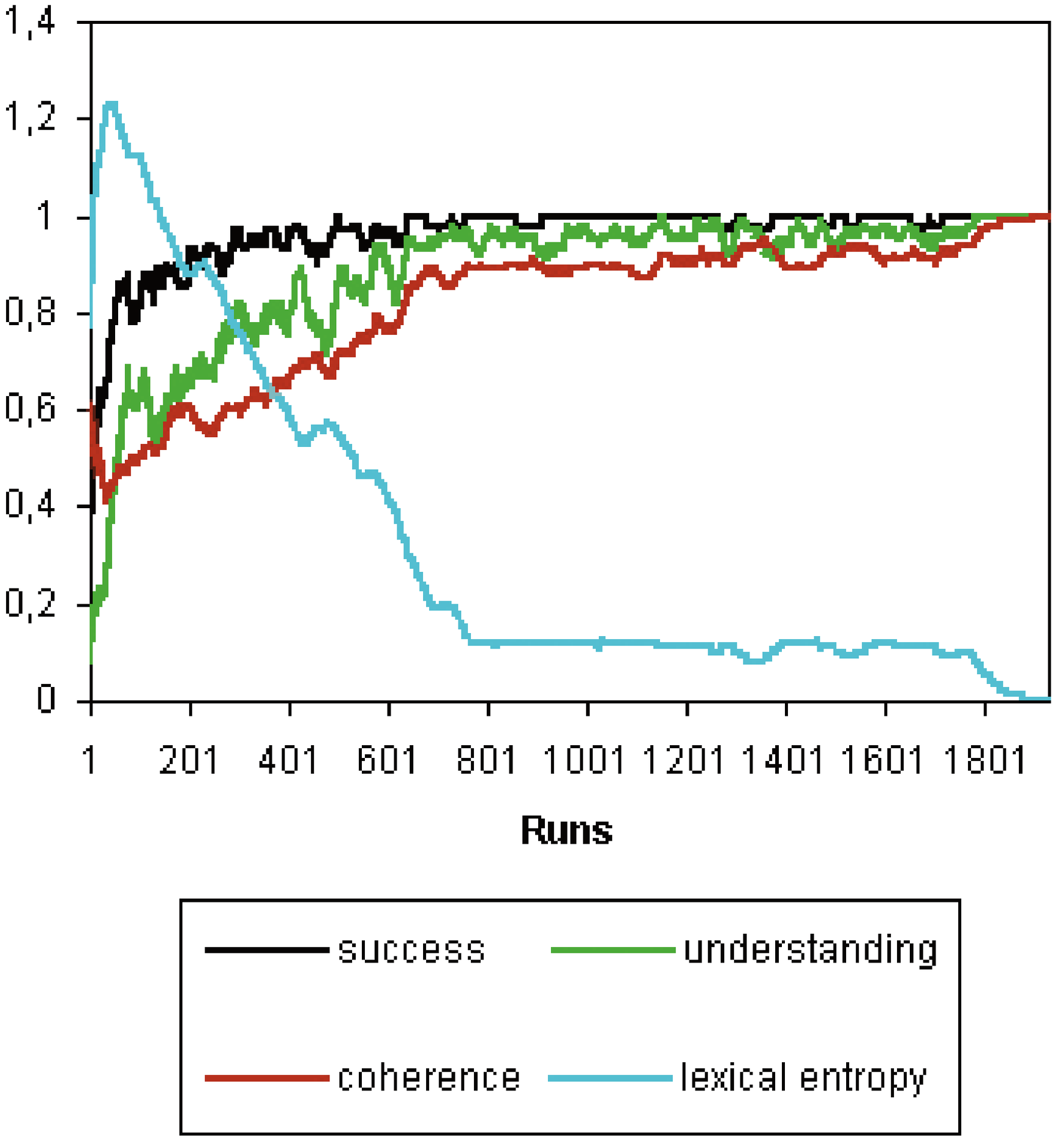
Look into the box explored how a population of artificial agents could invent and share a set of color distinctions and names to describe them. It therefore poses the fundamental question of whether machines can "create" new, perceptually grounded meanings and share them with each other. The names that the agents used for colors are not imposed by us and neither are their color distinctions. Nevertheless, we can observe in the experiment that agents gradually come to share the ways they categorize colors and the ways they name them. How is this possible?

The picture of the eye is first "pixelated" into small areas, and the three main colors of these pixels are chosen for further description. We see this in fig. 1, which shows a sample eye, the pixelated eye, and the three dominating colors for that eye. The agents were programmed to represent color stimuli in terms of color spaces, which is similar to the way humans are thought to represent color stimuli. One space is based on the brightness dimension (dark versus light), another space on the hue dimensions (red-green and yellow-blue), and still other spaces may be formed from combinations or subspaces of the previous two.

Making color distinctions boils down to dividing the color space into regions around prototypical points that act as the best example of a color. Thus, there could be a region around a prototypical blue and another one around a yellowish green. Agents can then invent their own distinctions by projecting new points in the color space, and they can categorize a perceived stimulus by calculating the distance using the existing prototypical points to find the one that is closest. For example, if a dark blue color is perceived, its stimulus values will be closer to the prototype for blue than to that for yellowish green. Agents can invent their own names by associating random strings, such as "ko" and "lecefi," with the prototypical points in their memory. And so when they see a color, they first compare it to the prototypes in their color spaces and then look up the name associated with the prototype that fits best. After a while the typical colors that appear in human eyes will have been recorded and the agents have a fairly stable repertoire of color categories and names to use for describing what they "see."

But if each agent creates its own color distinctions and color names, how can the group ever arrive at a shared set of categories and names? This is essential if agents are to have success in communication. The experiment explores the hypothesis that agents will align their distinctions and vocabularies. Recall that the agents play a language game about each color sample: One agent plays the role of speaker, categorizes the color, and produces one or more names. The other agent plays the role of hearer, looks up the names, and then compares their meanings with the name and categorization that it would have used itself. If there is not yet a name, the hearer could adopt the name used by the speaker. If there is already a name, the hearer could align its prototype to that of the speaker and increase the strength of the words used by the speaker in its own vocabulary. The result is that with these simple coordination dynamics, agents gradually align their color categories, and their conversations become more "coherent" as the speaker and hearer agree about the names used.

We have been doing this sort of color categorization and naming experiment for a while, studying carefully the conditions under which


1,4
1,2
1
0,8
0,6
0,4
0,2
0
1
201
401
601
801
1001
1201
1401
1601
1801
Runs
success
understanding
coherence
lexical entropy