



Universiteit
Leiden
The Netherlands

From monsters to mediators: The evolution of the theme of altruism in early robotic science fiction texts

Siglé, J.A.

Citation

Siglé, J. A. (2021, January 28). *From monsters to mediators: The evolution of the theme of altruism in early robotic science fiction texts*. Retrieved from <https://hdl.handle.net/1887/3134626>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3134626>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <https://hdl.handle.net/1887/3134626> holds various files of this Leiden University dissertation.

Author: Siglé, J.A.

Title: From monsters to mediators: The evolution of the theme of altruism in early robotic science fiction texts

Issue Date: 2021-01-28

Chapter 1: Science Fiction and Robots

Robots, Androids and Cyborgs

This chapter examines the differences between robots, androids and cyborgs in the context of science fiction before turning to relevant discourses on science fiction and robots more generally. According to Brian Stableford, the word “robot” was coined by Karel Čapek’s brother Joseph to mean forced labor (442).⁶ However, Stableford explains that a more contemporary definition came about when the “term was subsequently applied in a looser sense to industrial machines substituting for human laborers on automated production lines [...] and to machines capable of self-determined locomotion” (442). There is a certain degree of semantic overlap with its older equivalent, namely the automaton. Kang explains that an automaton can denote three different things, namely a self-moving machine (such as a wristwatch), a machine specifically designed to “mimic a living creature”, or a “person who acts like a machine in some way” (7-8). Kang also points out that “In the twentieth century, the word [automaton] has been eclipsed to a great extent by the modern ‘robot’” (8). In the context of science fiction, the *Oxford English Dictionary* defines a robot as “An intelligent artificial being typically made of metal and resembling in some way a human or other animal” (OED). Such a definition is sufficiently specific while not being overly prescriptive, and applies to all the artificial characters examined in this dissertation with the exception of the Steam Man, given that the Steam Man is not intelligent. However, the Steam Man character qualifies as a “mecha” (often shortened to “mech”), a term borrowed from Japanese anime and manga to mean: “a large armored robot, typically piloted by a person or creature inside the robot itself” (OED). According to these definitions, the Steam Man is a kind of robot that is piloted by a human (also discussed in Chapter 7).

Stableford explains that the origins of the word android can be traced back to “alchemical literature [...] with rumored attempts to create ‘homunculi’ by such alleged practitioners as Albertus Magnus and Paracelsus” (22). In science fiction, according to Stableford, androids are differentiated

⁶ Also discussed in Chapter 8 of this dissertation.

from robots because they have “synthetic flesh rather than inorganic components” (22). However, Stableford is quick to point out that “The usage is not consistent” (22). The *Oxford English Dictionary* similarly posits a vague and short definition of an android: “An automaton resembling a human being” (OED). Both definitions imply that the constitution of an android is less important, so long as it resembles a human. The third category, cyborg, is a “contraction of ‘cybernetic organism’, contrived to describe products of organic/inorganic chimerisation, particularly the augmentation of the human body with mechanical devices” (Stableford 114). According to the *Oxford English Dictionary*, a cyborg is defined as “A person whose physical tolerances or capabilities are extended beyond normal human limitations by a machine or other external agency that modifies the body’s functioning; an integrated man-machine system” (OED). Both definitions allow for an innumerable amount of possible combinations of different assemblages, while remaining specific enough to distinguish cyborgs from robots and androids.

In general terms, one can say that a robot refers to mechanical beings, as the term shares a certain semantic overlap with automaton as its predecessor. Robots may resemble humans, but do not have to. For example, the robots in the Hollywood adaptation *I, Robot* (2004) consists of humanoid machines that resemble humans, whereas *Wall-E* (2008) is also a robot but does not resemble humans. Abraham Merritt’s *The Metal Monster* (1920) contains a variety of robotic creatures not resembling human forms, while John Wyndham’s “The Lost Machine” (1932) is a first-person narrative about a robot that is described as a coffin with eight legs. Androids imply synthetic humans or artificial beings specifically designed to look like humans; their exact constitutions are less important. The most famous example is arguably Philip K. Dick’s *Do Androids Dream of Electric Sheep* (1968), in which the Voigt-Kampff test (measuring empathetic responses) is employed as a means with which to distinguish humans from replicants (with questionable accuracy). Cyborgs may or may not resemble humans, but must consist of assemblages of organic and inorganic materials. For example, the character of Robocop in *RoboCop* (1987) is part human (a brain and vital organs) and part machine (a mechanical chassis); the character of Will Carter played by Johnny Depp in

Transcendence (2014) downloads his consciousness into a computer, and can also be considered a cyborg as a result. While these terms remain popular, this dissertation considers them problematic when it comes to performing literary analyses of such artificial characters. For example, in *R.U.R.*, robots are made of synthetic materials mixed in large vats while remaining indistinguishable from humans in appearance, yet they are famous for being *robots* and not androids (as stated in the title of the play). The prefix robo- in *Robocop* suggests that it is a robot, while the character is more accurately a cyborg. All three terms may be applied to the T-800 character from the *Terminator* franchise. The T-800 refers to itself as a “cybernetic organism” (or cyborg) in the second film (1991), given that it consists of a robotic endoskeleton covered with synthetic skin. While the T-800 infiltrates 1980s and 1990s American society, it dons its synthetic skin and is indistinguishable from humans, qualifying it as an android. In glimpses of the postapocalyptic future, the T-800 is shown without synthetic skin because there is no longer a need to infiltrate human society, thus making it entirely robotic. Such transformations, however, have no bearing on the character which remains stable and uniform. Conversely, the T-800 evolves from being an antagonist to a protagonist in the first two films, but does so without having to change its constitution as a cyborg.

In other words, when analyzing such characters, these designations fail to reveal anything informative about their functions and significances within their respective narratives. The argument is not that these terms are entirely unusable or uninformative. One can certainly surmise in general terms that androids, by virtue of being doppelgängers, are more likely to incorporate themes regarding identity formation; cyborgs, by virtue of enhancing or altering the human condition via technology, are more likely to incorporate themes regarding technological monstrosities; robots, by virtue of emulating humans through artifice, are more likely to incorporate themes regarding free will and oppression. However, in the context of this dissertation, such generalizations are not particularly informative given that any such association is not a matter of veracity but convention. Whether dealing with a cyborg, robot or android, it is important to take each character in turn and not presume anything on the basis of these categories. Given the selection of texts in Part II, the

term robot (and robotic) remains the most appropriate to describe these characters in general terms, and where relevant, further clarification is provided.

Kathleen Richardson, in her book *An Anthropology of Robots and AI* (2015), discusses the role of anthropomorphism in relation to real robots, especially in the context of developing social robots which are designed specifically for the purposes of interacting (socially) with humans. Richardson explains that “anthropocentrism is not merely a frame within which to understand human interactions with technologies that have human qualities, but is also a means by which to rethink the importance of human sociality in underscoring interactions with nonhumans” (74). In the pursuit of social robots, questions regarding the role of anthropomorphizing machines is an important one as it determines the effectiveness with which robots may one day integrate into society. In their article, “Anthropomorphism as a Function of Robot Group Membership” (2012), Friederike Eyssel and Dieta Kuchenbrandt discovered in their experiments that “participants rated the in-group robot more favorably and anthropomorphized it more strongly than the out-group robot. That is, given the same background information about the robot (i.e. a picture), the mere manipulation of group membership affected subsequent evaluations of the robot” (729). The more the robot seems to belong to the same group as humans, the more likely humans tend to anthropomorphize them. While this is true in social robotics research, this dissertation is more focused on how groups are established in early science fiction narratives concerning robots. The term ‘group’ can denote any loose aggregate of people based on any perceived similarities as exhibited by the characters in the narrative. Although anthropomorphism plays an important role for the development of social robotics, this dissertation agrees with Gary K. Wolfe’s realization that the question of anthropomorphizing machines has been present since the first science fiction robot narratives. As Wolfe argues in “Icon of the Robot” (1979), in relation to science fiction robots:

This implied mythology is evident in even the earliest robot stories of modern science fiction [...] once a simple function of human being has been replaced, where will it stop? And if a

single human function can be supplanted by a mechanism, is it not possible that all human functions might one day be so replaced? (155)

If the implied mythology is to eventually replace all human functions, then social robotics is hard at work to fulfill this very mythology – not merely by replacing specific human functions, but also replacing human-human interactions as well. The ways in which science fiction deals with robots in turn define how humans come to see themselves, and the notion of social robots is merely a continuation of this very old process. As Wolfe argues:

Technology not only creates new environments for humanity, it also creates new images of humanity itself, which tend to mediate between the natural environment of mankind and the artificial ones it has created, between the past and the future, and between the known and the unknown. (151)

The Singularity would signify the ultimate amalgamation of humans and machines, even if by means of extermination.

Science Fiction and Robots

In *Science Fiction and the Moral Imagination* (2017), Russell Blackford briefly discusses the overlap between science fiction and posthumanist discourses. Many science fiction authors, according to Blackford, indirectly engage with posthumanist considerations, such as “the demise of human exceptionalism” and “the future of humanity as a species” (179). He points out that science fiction authors “were doing this before posthumanism and transhumanism were heard of” (179). What is relevant for this dissertation is Blackford’s observation that while transhumanism originated “in something like its current form during the 1980s”, one should not neglect “proto-transhumanist ideas expressed [...] early in the twentieth century” (179). Blackford’s analysis of science fiction is premised on its relation to what he calls “the moral imagination” (14). This is not to be confused with philosophies of morality or ethics, as he explains, but rather to make the claim that science fiction can be studied by examining its “contributions to humanity’s moral imagination”, and thereby “examine SF’s engagement with moral questions mainly for its own sake” (15). In this context

outlined by Blackford, the question of science fiction robots and altruism becomes all the more pertinent.

Approaching the emergence of science fiction from a different perspective, John Rieder in *Colonialism and the Emergence of Science Fiction* (2008) provides many innovative readings of early science fiction narratives. According to Rieder, the nature of humankind is central to any body of literature, but “scientific accounts of humanity’s origins and its possible or probable futures are especially basic to science fiction” (2). He examines the “complex mixture” of Darwinian ideologies concerning “competition, adaptation, race and destiny” as prevalent “thematic material of early science fiction” (2). It is no surprise, then, that science fiction originated in countries “most heavily involved in imperialist projects – France and England” and later, “the United States, Germany and Russia” (3). Rieder is careful to point out that his fundamental argument is not that “colonialism is science fiction’s hidden truth”, but rather: “science fiction exposes something that colonialism imposes” (15). Through science fiction’s complex treatments of such themes, it is important to “determine how and to what extent the stories engage colonialism” (3). Given Rieder’s insights, one must distinguish between early science fiction and texts involving fictional robots. While Rieder identifies a link between the emergence of science fiction and colonialism, which is likely to play a part in robotic science fiction during this era, robots as cultural and fictional objects also precede the history of science fiction. Examining early robot narratives, then, introduces a complex mixture of meanings in a vast network which cannot be fully elucidated in a single dissertation. While this dissertation therefore does not deny the expansive histories of both science fiction and robots, it must by necessity narrow its focus onto a specific thematic concern which, as mentioned, is the various engagements with themes of altruism.

When analyzing fictional robots more generally, some scholars adopt a transhistorical perspective, spanning from antiquity to Hollywood films. Such transhistorical approaches include Eric G. Wilson’s *The Melancholy Android: On the Psychology of Sacred Machines* (2006); Minsoo Kang’s *Sublime Dreams of Living Machines: The Automaton in the European Imagination* (2011); and

Despina Kakoudaki's *The Anatomy of a Robot: Literature, Cinema, and the Cultural Work of Artificial People* (2014). Other approaches that may be described as more synchronic include a much larger body of science fiction scholarship, such as Patricia S. Warrick's *The Cybernetic Imagination in Science Fiction* (1980), focusing particularly on robotic science fiction from 1930 until 1977; and Adrienne Mayor's *Gods and Robots: Myths, Machines and Ancient Dreams of Technology* (2018) which focusses on transcultural myths of artificial humans in antiquity.

Warrick's analysis, ranging from 1930 to 1977, covers as many as "225 short stories and novels" specifically related to robots and computers (xv). For Warrick, cybernetics "comprises all systems, mechanical and biological, in which information plays a role", and "is concerned with governance or control in social systems and in mechanical systems" (9).⁷ She identifies three kinds of systems, namely isolated, closed and open, and subsequently categorizes her selection of science fiction narratives in relation to each of these categories. The closed-system model is a reductionist model, concerned with mass and energy, and predicated on classical mechanics (Warrick 97); many of these narratives can also be described as dystopian. In contrast, the open-system model is closer to biological or living systems, "steady-state systems that exist in nonequilibrium", and capable of transformation that always involves "moving toward increased complexity" (Warrick 98). The isolated-system model is an ideal that does not exist in the real world, and consists of an "assemblage or combination of things that is uninfluenced by anything outside it" (Warrick 98-9). Warrick posits that traditional literary devices (plot, character, setting, etc.) are "secondary to other elements that are the central concerns of SF" (81). The central concerns of science fiction, according

⁷ It is worth mentioning that Warrick also discusses J. C. Maxwell's paper presented in 1868 to the Royal Society discussing the use of "governors" on steam engines, and links the etymological root of the Latin *gubernator* to the Greek *kybernētēs* (9). Hayles also retraces the history between governors and cybernetics when explaining the role of feedback loops in the early developments of cybernetics: "Feedback loops had long been exploited to increase the stability of mechanical systems, reaching a high level of development during the mid-to-late nineteenth century with the growing sophistication of steam engines and their accompanying control devices, such as governors" (8). Norbert Wiener, in his book *The Human Use of Human Beings* (1954), explained his motivation for employing the term cybernetics: "Hence, 'Cybernetics', which I derived from the Greek word *kubernētēs*, or 'steersman', the same Greek word from which we eventually derive our word 'governor'" (15).

to Warrick, can be identified as five elements that determine the strength of a science fiction narrative, namely scientific knowledge, novelty, dislocation in space or time, an “*awareness of unity* [original emphasis]”, and an emphasis on the mind (as a genre of ideas) (82-4). With these five elements successfully combined in a narrative, a sixth criterion measures the success of the overall narrative, namely that the reader experiences a new awareness, surpassing one’s “previous perceptions of time and space” (Warrick 84). Although such criteria are perhaps overly prescriptive for a general approach, Warrick’s analysis of robots and computers highlights the importance of the relationship between science and fiction: “We need to remind ourselves constantly that science and poetry are not separate universes; the working of the scientific imagination and the literary imagination parallel each other” (235). Thus, science fiction must “grow with science” and not oppose it: “it cannot react against science and still be *science* fiction” (237). Such a union between “the literary imagination that immerses itself in science” can “lead us, intelligently and humanely, into our future” (237).

Eric G. Wilson’s analysis also employs a transhistorical approach from a psychoanalytic perspective. Connecting the practice of mummification in ancient Egypt, the figure of the golem and various representations of automata, Wilson’s focus is primarily on the relationship between creators and their creations: “the androids emerging from human imagination constitute psychic projections as much as physical collections” (2). The practice of creating artificial doubles represents the “physical activities (the descent into the unconscious)” which “are analogous to spiritual activities (the ascent to soul)”, and despite their opposition, “they reach the same end” (4). In both cases, one finds the recursive movements, “from visible to invisible, outside to inside, known to mysterious” (4). Thus, to “make an android” is to “walk this razor’s edge between transcendence and neurosis” (4). His discussions ultimately conclude open-endedly:

From these tensions stems the nervousness most feel in this hyper-technological technological age. Those who love an organic cosmos suffer sporadic guilt from their repressed affection for their machines. Those who laud a mechanistic universe suspect they

are divorcing themselves from the vitalities that haunt their dreams. Both types attempt to suppress this chronic disorientation. But to repress is to grant control. [...] Both nature lover and computer maven risk becoming somnambulists: organic machines, mechanistic organs, living systems unaware of the forces driving them. (133-4)

Wilson argues that by embracing these figures, bringing them “into the playful light of the mind, to examine alike their accursed qualities and sacred potentials” will “force us to acknowledge seemingly inhuman beings [...] as our own” (139). Wilson’s analyses are particularly poignant at examining the very human psychological dimensions and tensions in artificial characters.

Minsoo Kang’s transhistorical approach also highlights the various positive and negative reactions to automata throughout different historical periods. His analysis similarly reflects on the inherent duality of automata as transgressors of preestablished oppositions:

[...] for a full understanding of the automaton motif in the Western imagination as a whole, one must take into account both aspects of the mechanical entity, to see how the object functioned in different historical contexts as the representation of *both* human empowerment *and* oppression, liberation *and* subjugation, transcendence *and* debasement. (305)

As Kang summarizes, during the medieval and Renaissance periods, the automaton reflected “the danger of demonic agency” and/or “the marvelous potential of natural magic”; during the Enlightenment, it became stripped of its “magical aura” and reflected “the rational order of the world, the state, and the body”; during the late Enlightenment, it reflected “a shabby representation of people lacking freedom either through oppression or conformism”; during the Romantic era, the uncertainty of the times meant the arrival of the “uncanny automaton appearing in literature”; during the industrial era, machines were “portrayed as living superhuman creatures”, which became “significantly darker and cautionary in the period after World War I” (304-5). Kang emphasizes that this ability “to hold such disparate meanings” is what makes the automaton “such a powerful and enduring conceptual object” (305).

Despina Kakoudaki in *Anatomy of a Robot: Literature, Cinema, and the Cultural Work of Artificial People* (2014) adopts a transhistorical approach, examining robots from antiquity to contemporary Hollywood cinema. She employs four “networks of meaning” spanning two millennia of artificial people; these networks include “artificial birth”, the “mechanical body”, artificial slaves and “the interpretation of artificiality as an existentialist trope” (26). These four elements can be found throughout historical and modern narratives of artificial people, and serve to reveal how “ancient ideas of animation and transformation inform modern legal and political narratives of objectification and subjectification, while the mechanical understanding of the body becomes associated with questions of control, agency, and psychological interiority” (212-3). Tracing the modes of thought throughout various historical periods, Kakoudaki concludes that the “contemporary theories of post-humanism” which seek to “redefine the human in a wider, more embodied, and more ecological frame of reference” must “counter nostalgic or apocalyptic trends that promote new body-phobic dualisms, notions of high tech transcendentalism and cerebral supremacy” (218-9). Kakoudaki convincingly demonstrates that artificial people have functioned, for a long time, as tropes with which to navigate “intimate philosophical questions”, such as: “What does it mean to be alive? [...] And how we can regard and treat one another without the prospect or threat of objectification?” (28). This latter question, in the context of this dissertation, is similarly preoccupied with ideas of altruism.

While these examples of artificial people and robotic scholarship inform the underlying perspective of this dissertation, it is also important to point out the differences in intention and methodology. Given the expansiveness of such transhistorical approaches, one rarely encounters close readings of specific texts. As Kakoudaki notices in relation to Hayles and Haraway, such “philosophical interpretations tend to reference literary and cinematic texts but quickly move on” (8), whilst later acknowledging herself that: “I use quick readings of well-known texts and films in order to ground complex philosophical or transhistorical questions” (27). While the theoretical authors discussed thus far have contributed invaluable research towards the broader aim of

understanding the various modes and semiotic structures of meaning of automata throughout history, this dissertation makes a far more localized and specific contribution to these considerations, one that Warrick, Wilson, Kakoudaki and Kang do not engage with, namely the ways in which early robotic science fiction deals with themes of altruism. Such a contribution necessitates close readings as opposed to “quick readings” of the selected texts. Whereas Warrick deemphasized the importance of traditional literary devices, this dissertation approaches robotic science fiction from the opposite direction: the various manners in which altruism features in these texts only become evident when focusing on characters’ behaviors, sentiments, motivations and interrelationships with other characters.

It should also be explained that, as many of these authors similarly explain (Warrick xiv; Kakoudaki 27), the prevalence of automata throughout history in the Western imagination inevitably means the elimination of many science fiction narratives for the sake of scope. This dissertation similarly selected five well-established works of robotic science fiction in its early years, namely *Frankenstein* (1818), *The Steam Man of the Prairies; or The Huge Hunter* (1868), *Tomorrow’s Eve* (1886), *R.U.R.* (1921), and the novel (not the film) *Metropolis* (1927). This selection is made on the basis of the various idiosyncratic treatments of altruism by each narrative. This selection does not suggest that other texts from the same period fail to incorporate this thematic dimension, nor to assert that all such narratives include these themes, but merely that this particular selection of well-known early science fiction robots already serves to demonstrate the fundamental claim that early science fiction robots engaged with themes of altruism in idiosyncratic ways.

Examples of science fiction narratives that were omitted include texts such as Edgar Allan Poe’s “The Man That Was Used Up” (1839), in which a general, consisting merely of a “large and exceedingly odd looking bundle of something”, is physiologically reassembled with prosthetics (40). Herman Melville’s “The Bell Tower” (1856) includes an unfortunate accident when the protagonist is killed by a mechanical human figure. Ambrose Bierce’s “Moxon’s Master” (1899) is about an inventor who is killed by his own chess-playing machine. Although one can speculate whether this is

an early warning of artificial intelligence, the reliability of the narrator casts doubt as to the validity of such an interpretation: "This sequence of events is usually interpreted as a story about a vicious robot that murdered its inventor, but it is much more likely to be a murder mystery in which the young man was skillfully misled by the actors" (Bleiler 65). George Haven Putnam's *The Artificial Mother; A Marital Fantasy* (1894) is about a mechanical babysitter who competes with the biological mother, although the entire event is revealed to be a dream. E. V. Odle's *The Clockwork Man* (1923) features a cyborg amalgamation of a human and clockwork mechanics, sent from the future to warn humanity not to create the very future from whence it came. Gaston Leroux's *The Bloody Doll* (1924) is a precursor to the Hollywood film *Child's Play* (1988). In *The Bloody Doll*, the brain of a recently executed criminal is transplanted into a machine (or puppet) and subsequently goes on a killing spree. S. Fowler Wright's "Automata" (1929) is an early depiction of "the machine as the next stage in the evolutionary process" (Warrick 51), although concerning the loss of humanity at the hands of machines, does not portray explicit scenes of altruism (Wright). E. M. Forster's *The Machine Stops* (1909) certainly deals with humanity's transformation into a hive (as the opening sentence demonstrates), as well as exhibiting the augmentation of humanity and the disintegration of family relations when mediated by machines, yet it does not feature any automata. In "The Clericomotor," published anonymously in the Detroit Free Press newspaper in 1884, a robot gives a sermon to a congregation, until the boy responsible for turning the crank starts to turn it backwards, causing the robot to malfunction, resulting in the congregation to flee in horror ("The Clericomotor"). The December 1899 issue of *The Black Cat* magazine featured the comical story by Elizabeth Whitfield Croom, entitled "Ely's Automatic Housemaid," about robots being employed for domestic housekeeping (Croom). Although humorously depicting a few mishaps, it ultimately seems optimistic about such socio-technological prospects. A follow-up story published in the same magazine in October 1900, entitled "Mr. Corndropper's Hired Man" by W. M. Stannard, features a robot from the same inventor as Croom's narrative, this time working on a farm (Stannard). Stannard's depiction is similarly very enthusiastic about the prospects of robots working in the farmlands. Clare Winger

Harris' "The Artificial Man" (1929), is about a man who, after a series of accidents, replaces his limbs and organs with artificial prostheses and organs, effectively turning him into one of the earlier representations of a cyborg.

As discussed in Chapter 6, this dissertation begins its investigation with *Frankenstein* as the first major robotic science fiction narrative, and stops at *Metropolis* (1927). Kang similarly stops his study in the interwar period and admits it "might seem arbitrary, but there are number of good reasons for concluding the narrative proper here" (297). After World War II, as Kang explains, there was a "veritable explosion" (297) of science fiction robots: "The sheer bulk of the material that would have to be covered for a comprehensive study of the robot symbolism in the postwar era would require another full volume" (298). This dissertation certainly agrees but would also emphasize that the explosion of robots started to occur prior to World War II, with the inception of the so-called Golden Age of science fiction in 1930 (Warrick xvii; Blackford 27; MacArthur 13). Warrick begins her study at 1930, and adopts a very different approach to deal with the vast amount of narratives available. As Sian MacArthur similarly observes: "In the 1930s science fiction reached what is now widely regarded as its 'Golden Age' whereby the 'rules' and expectations of the genre and its many sub-genres became more readily identifiable and easier to classify" (13). In the case of robotic science fiction, one thinks of Asimov's three rules of robotics as the pinnacle of formalized humanoid robots (also discussed in the conclusion of this dissertation).⁸ Since the 1930s, the sentimental robot meant that 'rules' became necessary with which to mediate human-machine

⁸ While the 1920s gave rise to *R.U.R.*, *The Clockwork Man* and *Metropolis*, the 1930s became preoccupied with sentimental robots. John Wyndham's "The Lost Machine" (1932), Lester del Rey's "Helen O'Loy" (1938), and Eando Binder's "I, Robot" (1939), are the first robotic narratives that explore the emotional dimensions of their characters. In "The Lost Machine", the first robot narrative in first-person, is about a lonely robot that eventually commits suicide; "Helen O'Loy" is about a robot that marries and, after the death of the robot's husband, asks the engineer to be shut down; Binder's "I, Robot", also in first-person, is about a robot that is wrongfully accused of killing its maker and is persecuted by an angry mob. After reading *Frankenstein*, the robot understands why it is persecuted but decides to commit suicide as an act of altruism rather than fight and possibly injure several humans. Both "Helen O'Loy" and "I, Robot" were inspirations for Asimov's first robot narrative "Robbie" (1939). According to Asimov, his first robot narrative, although not directly referencing the three laws of robotics, nevertheless reveals that he had the three laws already in mind: "Asimov [...] says that clearly he was already thinking about The Three Laws of Robotics since the story makes a rambling reference to the First Law" (Asimov "Robbie" 68-9).

relationships when machines acquired human emotional dimensions; as the robot character of Adam Link puts it: “Ironic, isn’t it, that I have the very feelings you are so sure I lack” (18). However interesting it would be to continue the approach outlined in this dissertation into this era of popular science fiction robots, such an analysis falls outside the scope of this dissertation. As it stands, this dissertation contributes to a large body of scholarship and discourses on both science fiction studies as well as various discourses on the nature of conceptual robots more generally and hopes to stimulate further research and discussion on the nature of, and relationships between, humans, robots and altruism.