

**THE BRAIN IS WIDER THAN THE SKY:  
ANALOGY, EMOTION, AND ALLEGORY**

*Paul Thagard*

*University of Waterloo*

[pthagard@uwaterloo.ca](mailto:pthagard@uwaterloo.ca)

*Draft 6, June, 2010*

**INTRODUCTION**

Literary allegories are based on underlying analogies, as in John Bunyan's *Pilgrim's Progress*, whose characters proceed on a journey analogous to the paths of ordinary souls toward heaven. Hence we should be able to get a better understanding of how allegories work by applying cognitive theories of analogical thinking. This article is in part an application to allegory of the multiconstraint theory of analogy developed by Holyoak and Thagard (1995), but it goes beyond that theory by considering neurological as well as psychological processes. I am inspired by Emily Dickinson's evocative poem that begins:

The brain is wider than the sky,  
For, put them side by side,  
The one the other will include  
With ease, and you beside.

I interpret this stanza as pointing to the enormous representational power of the human brain that enables it to portray, not only the universe, but also the person who does the representing. I will describe how new theories of neural representation, encompassing both cognitive and emotional aspects, have the potential to help make sense of many kinds of literary comparisons, from poetic metaphors to parodies to literary allegories.

The main text analyzed in this paper is George Orwell's *Animal Farm*, whose effectiveness I will explore using the multiconstraint theory of analogy supplemented

June 9, 2010

with observations about neural functioning. I will analyze the structure of Orwell's allegory using *cognitive-affective diagrams* that illuminate how he uses the fable about farm animals rebelling against their owners to castigate what he saw as the Stalinist betrayal of the ideals of the Russian revolution. The effectiveness of Orwell's cognitive and emotional representations depend on underlying neural mechanisms for comparisons and evaluations. Putting these processes together generates a cognitive-affective theory of allegory that not only explains successes such as *Animal Farm* but also less compelling exercises such as José Saramago's *Blindness*. Before diving into literary matters, however, I need to provide a quick review of the relevant cognitive and neural theories.

### **ANALOGY, EMOTION, AND BRAINS**

To grasp how allegories work in the human mind, we need to understand the workings of analogy, emotions, and the brain in general. What follows is only a brief sketch, reviewing ideas that are developed much more fully elsewhere (e.g. in Gentner, Holyoak, and Kokinov, 2001; Holyoak and Thagard, 1995; Thagard, 2000, 2005, 2006, 2010).

Analogies are systematic comparisons in which a source situation provides information about a target situation. In allegories, the source is the literary work created by the writer and the target is some aspect of the human condition; for example spirituality is a target for the source *Pilgrim's Progress* and politics is a target for the source *Animal Farm*. Using the literary source to illuminate the real-life target requires people to generate a mapping between the source and the target that reveals what aspects (characters, properties, relations, events, etc.) of the target correspond to what aspects of

the source. Readers of *Pilgrim's Progress* need to appreciate that its main character, Christian, represents an ordinary person in real life - an everyman.

From the perspective of cognitive processing, analogical mapping is a difficult feat, because there are a great many possible ways in which two complex situations can be put into correspondence with each other. Psychological evidence and computer models suggest that our minds handle this complexity thanks to three constraints that help to direct analogical mapping: similarity, structure, and purpose. These constraints work together to ensure that the properties and relations in a source domain can usefully be transferred over to a target domain.

The similarity constraint is most obvious. An element of the source and an element of the target are more likely to be put into correspondence with each other to the extent that they are similar with respect to meaning or appearance. For example, it is easy to think of the character Christian as any ordinary person because they are both people and have all their standard traits and behaviors.

Structure is more complex, encouraging correspondences that systematically carry over relations from the source to the target. For example, the reader of *Pilgrim's Progress* needs to appreciate that Christian's travelling from the City of Destruction to the Celestial City has the same relational structure as people journeying from the ordinary world to heaven. Purpose concerns the cognitive function that the allegory is supposed to accomplish. In religious allegories, for instance, the primary function is inspirational and pedagogical, using an imagined example to teach people how to live. The purpose of political allegories may include both explaining aspects of a previous historical situation and providing lessons about how to avoid past mistakes. An additional purpose of an

allegory may be entertainment, amusing people to see correspondences between an imagined world and the real one.

Analogies often perform their functions by generating emotional reactions, such as awe and fear in religious contexts, inspiration and determination in political contexts, and amusement in entertainment contexts. Hence understanding of allegory requires a theory of emotions, which is a highly controversial topic in current cognitive science. Some theorists defend appraisal theories of emotions, taking them to be judgments that evaluate a situation with respect to how well it accomplishes goals. For example, an event in a novel will make the reader happy if the event involves something good happening to a character that the reader likes and thereby satisfies the reader's goal of wanting good things to happen to favored characters. In contrast, many theorists defend physiological theories of emotions that construe them as reactions to changes in bodily states such as heart rate and breathing patterns. However, these two approaches can be synthesized by a view of emotions as neural processes that integrate appraisal and physiological perception (Thagard and Aubie, 2008; Thagard, 2010). From this perspective, the emotions generated by allegories and other analogies are the result *both* of judgments that the reader makes about the allegorical source along with its real-life target *and* of physiological responses that result from reading or hearing the allegory.

If this integrative theory of emotions is correct, then emotions need to be understood as brain processes, not as abstract cognitive states. This development is in keeping with the increasing trend in cognitive psychology to understand mental processes as brain processes, not as computations independent of the underlying physical structure (see e.g. Smith and Kosslyn, 2007; Anderson, 2010). According to cognitive

neuroscience, we can think of mental representations such as concepts as patterns of firing activity in populations of neurons. A neural representation is not a static object like a word on paper or a street sign, but is rather a dynamic process involving ongoing change in many neurons and their interconnections. A population of neurons represents something by its pattern of firing.

The brain, with its 100 billion neurons, is capable of a huge number of patterns: if each neuron can fire 100 times per second, then the number of firing patterns of that duration is  $(2^{100})^{100000000000}$ , which is far larger than the number of elementary particles in the universe, only about  $10^{80}$ . I call this result “Dickenson’s theorem”, after the poem I quoted earlier: the brain really does have the capacity to represent the sky, the person, and the universe in general. A pattern of activation in the brain becomes a representation of something when there is a stable causal correlation between the firing of neurons in a neural population and what is represented, such as an object or group of objects in the world (Eliasmith, 2005; Thagard, forthcoming). The view that mental representations are patterns of firing in neural populations is a radical departure from everyday concepts, but is increasingly supported by experimental data and theoretical neuroscience.

In sum, the mental processes by which a reader represents, understands, and learns from an allegory depend on brain mechanisms for encoding information and using it to produce analogical inferences and emotional reactions. Let us now see how this view applies to one of the most influential modern allegories, Orwell’s (2008) *Animal Farm*.

## ANIMAL FARM

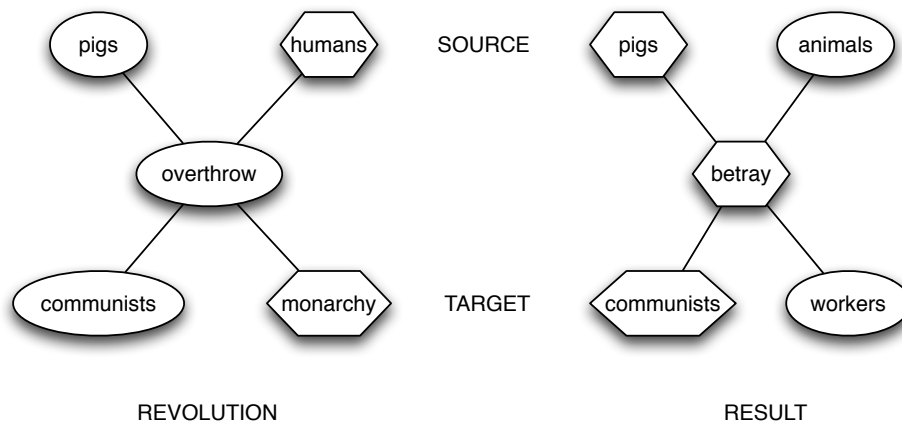
Since George Orwell published *Animal Farm* in 1945, it has been widely recognized as one of the great novels of the twentieth century, a scathing allegory for the development of communism in the Soviet Union. The book was based not only on Orwell's knowledge of Russian history but also on his conflicts with communists while fighting for the socialists in the Spanish Civil War. It is easy to construct the analogical mapping that Orwell used to highlight negative aspects of the Russian revolution. Here are some of the correspondences: the farm is Russia, the humans on the farm are aristocratic rulers of Russia before the revolution, the pigs are the communist revolutionaries, the other animals are the workers, the dogs are the secret police, Napoleon is Stalin (and possibly Lenin too), Snowball is Trotsky (and possibly other dissidents), the human Frederick is Hitler, and Boxer, the horse, is an ideal proletarian supporter of the communists.

Much more important than the correspondences of individual entities is the mapping of relations between those entities, including the causal relations between events. In *Animal Farm* the pigs overthrow the humans, just as in Russia the communists overthrew the aristocrats. Orwell clearly approves of both these initial events, but laments their results: the pigs end up betraying and oppressing the other animals, just as the communists ended up betraying and oppressing the workers.

Approval and lamentation are clearly emotional reactions, so the analogy that makes Orwell's allegory work has strong affective dimensions. These can be displayed using a new technique called *cognitive-affective diagrams* originally developed for analyzing political disputes (Findlay and Thagard, forthcoming). Concept diagrams

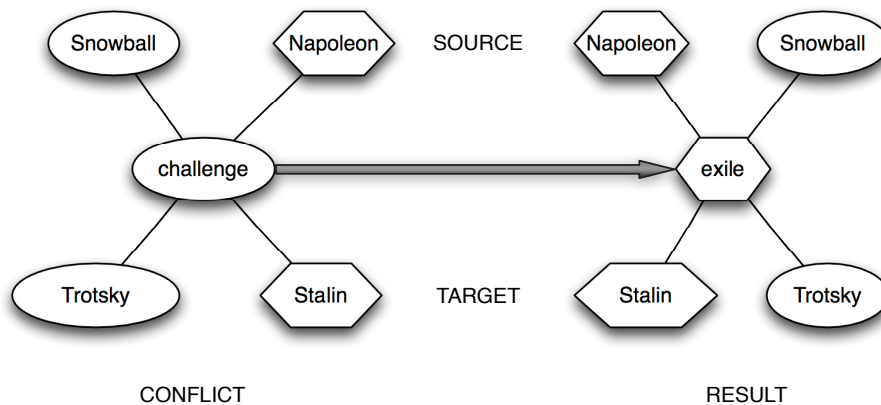
have been used for decades to analyze the structure of complex representations, but cognitive-affective diagrams add an additional dimension by indicating the emotional significance of the key concepts. (Concept diagrams are often called maps, but to avoid confusion with analogical mapping, this paper refers to cognitive-affective diagrams rather than maps.)

Figure 1 provides a diagram of some of the key relations in *Animal Farm*, using ovals to represent positive concepts and hexagons to represent negative ones. This figure shows how Orwell portrayed his fable as having not only the same relational structure as the Russian revolution, but also has having what he wanted to convey as the same emotional structure. Pigs and communists start out as emotionally positive, but become negative through betrayal of those whom they are supposed to represent.



**Figure 1.** Cognitive-affective diagram showing the analogical mapping between *Animal Farm* and the Russian revolution. Ovals represent positive concepts and hexagons represent negative ones. The source analog (the narrative) is shown on the top, and the target analog (history) is shown on the bottom, with relations in common.

Additional cognitive-affective diagramming can illustrate the broader range of correspondences between the emotional status of events in the book and events in history. For example, Snowball’s challenge to Napoleon leads to his exile. Figure 2 shows correspondences between the novel and the history with respect to individuals, relations between individuals, relations between events, and (just as important) between Orwell’s intended emotional connections between the two situations. It is crucial for the success of the allegory that the correspondence between the historical target and the allegorical source maintains emotional as well as analogical consistency, because otherwise the allegory fails to achieve its emotional purpose. Many additional aspects of the analogy could be illustrated by more detailed cognitive-affective diagrams, but I want instead to move on to say more about the underlying cognitive and neural processes needed to appreciate allegories.



**Figure 2.** Cognitive-affective diagram showing the relational and emotional correspondences between events in Orwell’s book and in the Soviet Union.

The three analogical constraints of similarity, structure, and purpose are all important to the effectiveness of Orwell’s allegory. There is not much visual similarity

between Orwell's characters and the leaders of the Russian revolution; for example, Napoleon is not described as looking at all like Stalin. Moreover, semantic similarity is somewhat confusing in the mapping between the farm and Russia, because all of the characters in Russia were human, whereas only some of the characters in *Animal Farm* are human (such as the prerevolutionary rulers); and the revolutionary pigs are not humans, although they often behave like them. However, some of the characters have other features that make them similar, for example the ruthlessness of both Napoleon and Stalin, and the cleverness of both Snowball and Trotsky. The long tradition of anthropomorphizing animals in fables for instructive purposes predisposes readers to notice such correspondences. The key relational concepts, moreover, are close to identical, as shown by the occurrence of *overthrow*, *betray*, *challenge* and *exile* in the depiction of both the fictional source and the historical target in figures 1 and 2.

Those figures also make clear the close and natural structural correspondences between *Animal Farm* and the Russian revolution. These correspondences roughly obey the two key aspects of the constraint of structure: 1-1 mapping and relational consistency. As allowed in the multiconstraint theory of Holyoak and Thagard (1996) but not in the structure mapping theory of Gentner (1983), 1-1 mapping may occasionally be violated, as in the possibilities that Napoleon seems to carry aspects not only of Stalin but also of Lenin. Snowball seems like a composite of Trotsky and other opponents of Stalin, and Boxer corresponds to many loyal workers rather than a distinct example. Still, this complication does not interfere with basic appreciation of the relational correspondence between Napoleon's exile of Snowball and Stalin's exile of Trotsky.

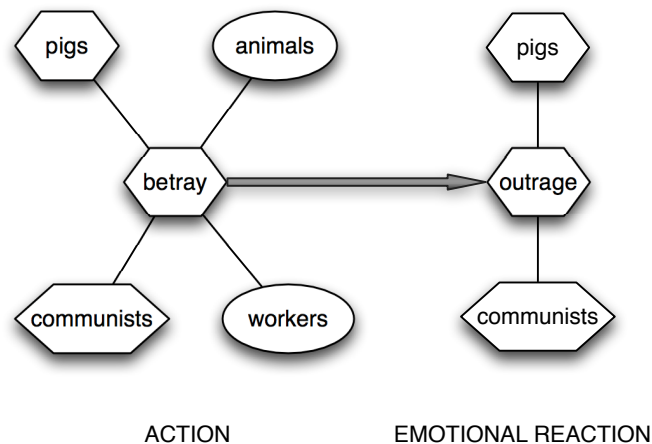
Relational consistency is maintained, in that the exile of Snowball by Napoleon maintains the ordering in the exile of Trotsky by Stalin.

Most importantly, the cognitive-affective diagrams in figures 1 and 2 clarify why *Animal Farm* is so effective at carrying out its allegorical function. Orwell's book was neither a barnyard fantasy designed to amuse children, nor a cold war diatribe against totalitarian Russia. It was written around 1943 at a time when the Soviet Union was allied with Britain and the United States in the war against Nazi Germany, and Orwell had difficulty finding a publisher that did not consider it unduly critical of Stalin's government. The book's purpose, in addition to entertainment and aesthetic achievement, was to provide a warning about the way in which the Communist Party could be ruthless about stamping out all opposition, as Orwell had observed in the Spanish Civil War. To this end, emotional correspondences shown in figures 1 and 2 are just as important as the relational ones, carrying the main message of the allegory. Napoleon is clearly a reprehensible figure in *Animal Farm*, and Orwell similarly wants readers to see Stalin as a ruthless and domineering, not (as he was commonly portrayed in 1943) the benign Uncle Joe mobilizing Russia against Hitler.

This emotional component of analogical mapping has been emphasized by various writers (Thagard, 2006, ch. 3; Blanchette and Dunbar, 2001; Bliznashki and Kokinov, 2009). Overt political arguments often have the purpose of transferring positive or (more often) negative attitudes from the source to the target analog. For example the American defeat in Vietnam and the Soviet defeat in Afghanistan are often used as emotionally negative source analogs to argue against current foreign involvement in Afghanistan. Political allegories are more subtle, leaving the analogical argument to

be generated by the reader, but the desired effect is emotionally similar, and may even be more powerful because of the reader's engagement in its completion. For *Animal Farm*, the effect on the reader would be something like this: Just as you are appalled by Napoleon's betrayal and exploitation of the animals who were supposed to have been liberated from human domination, so you should be appalled by Stalin's suppression of dissent and adoption of policies beneficial only to a ruling elite rather than the workers who were supposed to have been liberated from exploitation.

Figure 3 roughly shows how the outrage generated by the emotional reaction to *Animal Farm* can be transformed into an enhanced emotional reaction to the Soviet Union under Stalin. The purposes of political allegories, like political analogies in general, usually include transfer of emotional reactions from source to target. Orwell's later depiction of a totalitarian society in *1984* also provides a powerful source of emotional transfer to multiple targets including oppressive capitalist and fascist societies as well as the Soviet Union.



**Figure 3.** Cognitive-affective diagram showing how the emotion of outrage can be transferred from the allegorical source to the historical target.

Dystopian novels allegorically generate negative emotions to be analogically applied as warnings, while utopian novels can suggest possibilities intended to be attractive. For example, some of the science fiction novels of Ursula K. LeGuin explore possibilities for developing egalitarian, non-sexist societies. These utopias serve as sources that suggest ways of building new targets for future social developments.

I do not mean to suggest that the only purpose of allegorical works is to generate insights and arguments, as they can have additional literary and social functions such as providing entertainment. *Animal Farm* and *1984* are both fun to read even for the politically unsophisticated. Moreover, analogical function in literature can be diverse, even for parody, which typically serves to make fun of the target analog, but can also make fun of aspects of the source. For example, Weird Al Yankovic has a hilarious music video “White and Nerdy” that is funny not just because it satirizes gangster rap, but because it also makes fun of math geeks trying to be like rappers. LeGuin’s imaginary worlds are fascinating not only because they provide analogical suggestions of how our own society could be better, but also because they depict intrinsically interesting interpersonal relations.

### **HOW ALLEGORY FAILS**

I have suggested that *Animal Farm* is a wonderfully successful allegory because it satisfies all the constraints of analogical thinking while also producing a strong emotional effect. The constraints of similarity and structure are not perfectly satisfied, but the

mappings between the novel and the historical target are sufficiently clear to accomplish the emotional and political purposes. The allegory highlights approval of overthrow of oppressors, disappointment at betrayal of the oppressed, outrage at suppression of dissent, and indignation at distortions of history.

A much less successful allegory is *Blindness*, by the Portuguese author José Saramago (1999), who won the Nobel Prize for literature in 1998. The book is marketed (e.g. on the back jacket of the paperback edition) as “a magnificent parable of the loss and disorientation and a vivid evocation of the horrors of the twentieth century” and as a “surreal allegory”. It is a disturbing depiction of what happens when an epidemic of blindness hits an unnamed city, leading to social devastation. The two main characters are a doctor and his wife, also unnamed. Unlike many commentators who were enthusiastic about the book – it was a *New York Times* notable book of the year – I found it an unsatisfying concoction of ugly events, undeveloped characters, run-on sentences, and ill-expressed ideas.

Whatever its successes for some readers, the multiconstraint theory of analogy illuminates why *Blindness* is not an effective allegory. The epidemic of blindness is a potentially vivid source analog, but the book gives no clue to what are the target analogs. There are many grisly occurrences in *Blindness*, perhaps shadowing such historical atrocities as the Jewish holocaust and ethnic cleansing in the former Yugoslavia. But the main characters in the book, whose psychological states are rarely explained, bear no noticeable correspondence to any historical figures. The doctor’s wife does not succumb to blindness, but pretends to be blind in order to help her husband; but the significance of her exceptionality is opaque. None of the events in *Blindness* such as the incarceration of

the newly blind in an asylum can be put into correspondence to particular historical events. There are not even vague mappings available between the underspecified causal relations in the book and causal relations in historical atrocities. The book ends implausibly with unexplained cessation of the epidemic, another event which has no apparent correspondence to any historical event.

Because of these failures to satisfy constraints of similarity and structure, *Blindness* utterly fails to have any clear emotional or political purpose beyond a vague sense of dread. The brutality in the book definitely generates negative emotions such as disgust, and perseverance through adversity occasionally generates positive emotional reactions such as admiration. But despite its allegorical portent the book produces no instructive political lessons: to what contemporary situations should its emotions of outrage and despair be applied? Even a critic who thought that the book has intrinsic stylistic merits would have to concede that it fails as an exercise in analogy, and hence also, I would argue, as an exercise in the cognitive/affective process of allegory.

Some readers may find *Blindness* effective in other ways, for example in evoking anxiety, alienation, and dehumanization as Kafka's stories do much more effectively. If the book had more fully developed characters and a more plausible plot, then it could amount to a literary success according to valid criteria other than allegorical efficacy. But as a parable of modern times *Blindness* fails because of the inadequacy of the analogical mapping between the cast and events of the source novel and any recognizably outlined contemporary target domains.

## EMOTIONS AND ALLEGORY

For a deeper understanding of how allegories can accomplish emotional purposes, we need a fuller picture of the nature of emotions. My integrated theory of emotion uses neural mechanisms to show how emotions can be both judgments and physiological perceptions through the parallel interactions of diverse brain areas (Thagard and Aubie, 2008; Thagard, 2010). The relevant brain components range from prefrontal cortex areas capable of appraisal to subcortical areas such as the amygdala and insula that respond to bodily states. The brain is wider than the sky not only in being able to represent the observable and theoretical aspects of the world, but also in being able to appreciate their emotional significance.

Effective allegories like *Animal Farm* engage neural processes for both cognitive appraisal and physiological perception. Cognitive appraisal is engaged when the story activates the goals of the reader, such as social progress in accord with principles of justice. When the story reports the satisfaction of those goals, for example when the animals overthrow their human oppressors, then the reader feels happy. But when the story reports the blocking of those goals, for example when pigs distort the egalitarian principles of the revolution in the service of their own self-interest, then the reader feels negative emotions such as sadness and anger.

However, emotions are more than just cool judgments about the satisfaction of goals, for they also involve gut reactions or (to use the term of Damasio, 1994) somatic markers. Emotions are tied to physiological changes such as heart rate, breathing, skin conductance, blood pressure, cortisol levels, and so on. An exciting or upsetting book can literally set the heart racing. A purely physiological account of emotions is not

adequate to capture the subtlety and range of human emotions, but nevertheless bodily changes need to be included in a full account of what happens in our brains when we feel emotions. An integrated theory of emotion uses neural interconnections to incorporate both appraisal of situations with respect to goals and perception of bodily states, and is thereby able to account for both the contemplative and the visceral aspects of emotional reactions to allegories and other forms of literature.

An important question in current cognitive science is how the brain manages to bind together different kinds of representation. For example, the reader of *Animal Farm* simultaneously represents Napoleon as looking like a pig, talking like a human, and behaving in contemptible ways. Such representations combine verbal, visual, auditory, and emotional representations. Drawing on techniques developed by Eliasmith and Anderson (2003), Thagard and Stewart (forthcoming) propose a neurocomputational account of how multimodal representations, including emotional ones, can be combined. The central notion in their account of combination is *convolution*, which traditionally means “twisting together” but can also be understood metaphorically, mathematically, and neurologically.

Metaphorically, convolution can be applied to describe any kind of conceptual combination, from mundane examples like “blue shirt” to highly creative ones like “sound wave”, “natural selection”, and “wireless email”. This application is akin to the suggestive metaphor of blending that Fauconnier and Turner (2002) have used to describe many kinds of conceptual integration. Two concepts can be convolved to produce a surprising new concept that may turn out to be valuable and hence qualify as creative. More rigorously, convolution can also be understood mathematically as an operation that

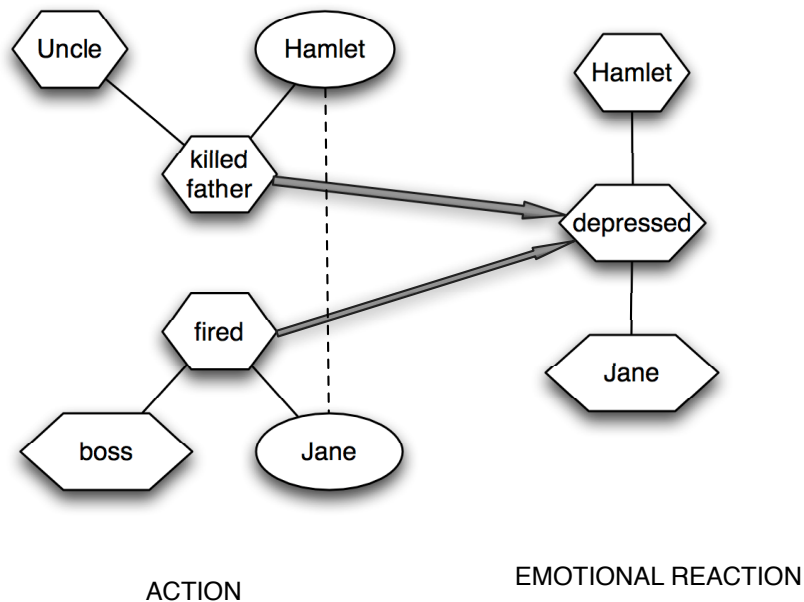
binds together wave functions or vectors (Plate, 2003), and this operation can be performed by neural populations (Eliasmith, 2005).

Thagard and Stewart (forthcoming) describe computational models of how concepts understood as patterns of activation in neural populations can be convolved by producing new patterns of activation that carry over much of the meaning of the original concepts. Moreover, the resulting new concept can receive an affective evaluation through convolution with patterns of neural activity that integrate both the appraisal and physiological aspects of emotion. The resulting overall convolution binds together, in a mathematically rigorous and neurologically plausible way, combined concepts and emotional responses. Literary theorists who find blending helpful to describe the psychological functions of allegory can appeal to this notion of convolution to provide a neurocomputational explanation of conceptual integration.

### **THE FUNCTIONS OF FICTION**

Neural models of emotion and analogy are also relevant to general questions about how non-allegorical fiction works. Fiction has many functions such as when people use their theory of mind to attribute mental states to fictional characters (Zunshine, 2006), and when people simulate the world by means of abstraction, simplification, and expression (Mar and Oatley, 2008). For many people, a key part of the enjoyment and edification that fiction brings derives from empathic identification with central characters. Readers of novels or watchers of plays and movies can feel such reactions as the disappointment of Chekhov's Irina, the anxiety of Shakespeare's Hamlet, and the fear of Winston Smith in Orwell's *1984*.

Empathy is a kind of analogy in which one makes a comparison between the condition of someone else and one's own current or past condition (Barnes and Thagard, 1997). For example, a reader (call her Jane) might see an analogy between her own life and the story of Hamlet, in that being fired by her boss caused her to be depressed and indecisive, just as the murder of Hamlet's father caused him to be depressed and indecisive. We can use a cognitive-affective diagram to capture the structural and emotional similarities between the two situations (figure 4). By noticing the correspondences between her own situation and Hamlet's, Jane may accomplish both the purpose of understanding her own life better and the literary purpose of becoming absorbed in the story of Hamlet.



**Figure 4.** Cognitive-affective diagram of empathic understanding in response to *Hamlet*. The dotted line shows the analogical connection between Jane and Hamlet.

It is important, however, not to understand empathy as a purely cognitive analogical operation tied to the appraisal process of emotion. In real life, empathy often involves a more visceral process of feeling someone's pain or other experience through a process of emotional contagion furthered by mirror neurons (Thagard, 2010). Such structures were first identified in monkeys who have neural populations that fire both when they move an arm and when they see another monkey move an arm (Iacoboni, 2008). Brain scanning experiments have found pain-related neural areas in humans that seem to respond immediately to the pain of others. The kind of empathy that arises from perception of the emotional states of others can therefore be more immediate than the more verbal kinds of comparison usually modeled in analogical mapping. This more direct, visceral kind of empathy does not undermine the more verbal sort shown in figure 4: people seem to be capable of both kinds of empathic mapping. Plays have an advantage over novels in that observing the body language of actors can stimulate a person's mirror neurons, without the more indirect use of the imagination that is required for novels.

In the appreciation of fiction, the more visceral kind of empathy may play a role as important as the more verbal kind. Readers who are absorbed in an engaging novel may feel some of the same underlying physiological responses that are being described in the character, as the readers imagine themselves to be undergoing similar experiences. For example, if I imagine myself driving fast in a car chase, this may engender feelings of exhilaration and fear. I can then project this emotion back onto the character I am reading about, even if the author is not explicit in attributing the emotion to the character.

## CONCLUSION

This paper has used Orwell's fable *Animal Farm* to illustrate the cognitive and emotional processes that can make allegories effective for many purposes, from education to entertainment to political agitation. I have used the multiconstraint theory of analogy to explain what makes the book such a successful allegory, in contrast to the much more feeble exercise of *Blindness*. In addition, I have shown the relevance to allegory of very recent developments in cognitive science, including a new theory of emotional consciousness that uses an integrated understanding of appraisal and physiological perception to explain why reading novels, seeing plays, watching movies, and even playing video games can be psychologically powerful for allegorical and other purposes. An even newer technique, cognitive-affective diagramming, has been used to display the structure of allegories in both their analogical and emotional aspects.

EMPATHICA is a software tool currently under development that facilitates production of these diagrams. Its primary purpose is aiding conflict resolution by encouraging disputants to identify the cognitive-affective differences that generate disagreements; but EMPATHICA may also turn out to be useful for analyzing literary texts, and possibly even for sketching out new characters and situations in the development of new novels, plays, or games.

The observations in this paper about the relevance of emotional consciousness and cognitive-affective diagrams to understanding allegory apply much more generally to a wide range of genres that use similar psychological processes. Genres that rely on systematic comparisons to generate emotional responses include many instances of parody, satire, and poetic metaphor. Frye (1963, p. 30) suggests that whenever you try to

explain what a poem means you are bound to turn it into an allegory to some extent. It is encouraging that literary theorists are beginning to draw on the resources of cognitive psychology and neuroscience to help understand aesthetic responses (e.g. Thrailkill, 2007; Massey, 2009). I have tried to show how emotional brain processes of cognitive appraisal and physiological perception unite with analogical thinking to enable people to appreciate the delights of allegory and other literary forms.

**Acknowledgements.** This research has been supported by the Natural Sciences and Engineering Research Council of Canada. Thanks to Randy Harris for very thorough and helpful comments on previous drafts.

## REFERENCES

- Anderson, J. R. (2010). *Cognitive psychology and its implications* (7th ed.). New York: Worth.
- Barnes, A., & Thagard, P. (1997). Empathy and analogy. *Dialogue: Canadian Philosophical Review*, 36, 705-720.
- Blanchette, I., & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory & Cognition*, 29, 730-735.
- Bliznashki, S., & Kokinov, B. (2009). Analogical transfer of emotions. In B. Kokinov, K. Holyoak & D. Gentner (Eds.), *New frontiers in analogy research* (pp. 45-53). Sofia: New Bulgarian University Press.
- Damasio, A. R. (1994). *Descartes' error*. New York: G. P. Putnam's Sons.
- Eliasmith, C. (2005). Neurosemantics and categories. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 1035-1054). Amsterdam: Elsevier.
- Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering: Computation, representation and dynamics in neurobiological systems*. Cambridge, MA: MIT Press.
- Fauconnier, G., & Turner, M. (2002). *The way we think*. New York: Basic Books.
- Findlay, S., & Thagard, P. (forthcoming). Emotional change in international negotiation: Analyzing the Camp David accords using cognitive-affective maps.
- Frye, N. (1963). *The educated imagination*. Montreal: CBC Enterprises.

- Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gentner, D., Holyoak, K. H., & Kokinov, B. K. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press/Bradford Books.
- Iacoboni, M. (2008). *Mirroring people: The new science of how we connect with others*. New York: Farrar, Straus and Giroux.
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, 3, 173-192.
- Massey, I. (2009). *The neural imagination: Aesthetic and neuroscientific approaches to the arts*. Austin: University of Texas Press.
- Orwell, G. (2008). *Animal farm*. London: Penguin.
- Plate, T. (2003). *Holographic reduced representations*. Stanford: CSLI.
- Saramago, J. (1999). *Blindness* (G. Pontiero, Trans.). Orlando: Harcourt.
- Smith, E. E., & Kosslyn, S. M. (2007). *Cognitive psychology: Mind and brain*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.
- Thagard, P. (2005). *Mind: Introduction to cognitive science* (2nd ed.). Cambridge, Mass.: MIT Press.
- Thagard, P. (2006). *Hot thought: Mechanisms and applications of emotional cognition*. Cambridge, MA: MIT Press.

- Thagard, P. (2010). *The brain and the meaning of life*. Princeton, NJ: Princeton University Press.
- Thagard, P. (forthcoming). How brains make mental models. In L. Magnani, W. Carnielli & C. Pizzi (Eds.), *Model-based reasoning in science and technology. Abduction, logic, and computational discovery*. Berlin: Springer.
- Thagard, P., & Aubie, B. (2008). Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. *Consciousness and Cognition*, *17*, 811-834.
- Thagard, P., & Stewart, T. C. (forthcoming). The Aha! experience: Creativity through emergent binding in neural networks. *under review*.
- Thrailkill, J. F. (2007). *Affecting fictions: Mind, body, and emotion in American literary realism*. Cambridge, MA: Harvard University Press.
- Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Columbus: Ohio State University Press.