Douglas F. Watt

Social Bonds and the Nature of Empathy

Considerations stemming from a basic taxonomy of emotion suggest that the creation of social bonds is a critical domain for affective neuroscience. A critical phenomenon within this group of processes promoting attachment is empathy, a process essential to mitigation of human suffering, and for both the creation and long term stability of social bonds. Models of empathy emerging from cognitive and affective neuroscience show widespread confusion about cognitive versus affective dimensions to empathy. Human empathy probably reflects admixtures of more primitive 'affective resonance' or contagion mechanisms, melded with developmentally later-arriving emotion identification, and theory of mind/perspective taking. From these considerations, a basic model of affective empathy is generated as a gated resonance induction of the internal distress of another creature, with an intrinsic motivation to relieve the distress. It is 'gated,' in that at least four classes of hypothesized variables determine intensity of an empathic response to the suffering of another. Differential predictions of this model vs. current ones, and future tests are proposed.

'Where there are two, one cannot be wretched and one not.'

— Euripides

Taxonomy Must Come First: A Possible Typology of Emotion

Jaak Panksepp and I have both argued recently (Panksepp and Watt, 2003) that basic emotion typologies or taxonomies are critical to defining the problem spaces of emotion, and that our classifications schemes, if conceptually flawed, seriously impair meaningful data collection and only compound an already substantial degree of conceptual confusion. One possible typology for emotion (Panksepp, 1998; Panksepp & Watt, 2003 for some further implications) might

Correspondence:
Douglas F. Watt, Ph.D., Clinic for Cognitive Disorders, Quincy Medical Center, Boston University, School of Medicine, Quincy, MA 02169, USA. *Email: dfwatt@brahmacom.com*

Journal of Consciousness Studies, **12**, No. 8–10, 2005, pp. 185–209

emphasize the following primary or prototype states: ***Fear, Rage, Lust, Separation Distress, Play/Social Affection, Nurturance***, and also possibly ***Disgust***. I omit '***Surprise***' from this typology, despite its appearance in many other classical typologies for emotion, as this is not a primary emotional state, more a 'cognitive or expectancy reset' mechanism involving cognitive dissonance. Additionally, the notion of a '***Seeking System***' (see Panksepp, 1998, for summary overview) centred in the VTA-to-lateral hypothalamic DA to mesolimbic-mesocortical projection tracts, falls into a special class of one, distinct from the other prototype emotional systems. Without the 'seeking system', there is no ability to instantiate *any* coherent emotional state, and in a direct sense all the prototype states simply disappear, along with virtually any motivated behaviour. This suggests that the ***Seeking System*** might provide a kind of 'central trunk line' in the brain for all emotion that the prototype emotional states may 'grab a hold of' to instantiate and organize a particular emotional state in the brain.

Even in highly cognitive creatures such as ourselves, those simple primes or prototypes remain central organizing paradigms for our more cognized emotions. Fundamental paradigms of emotion, like confrontation with a predator or powerful rival (fear), assault, or other threats to safety, free pursuit, territory or other resources (anger), the loss of mates, offspring and con-specifics (separation distress), and the comforts and joys of sex, attachment, and play all become increasingly translated and activated through complex cognitive operations that sit on the cognition-emotion border. Through human ontogenesis, these affective 'primes' are increasingly activated within, and even transformed by, this rich sea of human symbolic operations and complex meanings, as emotion is increasingly cognized. But even complex, highly cognized human emotions never lose their intrinsic grounding to these emotional primes, or else they cease to exist in the human organism. nor does emotion ever lose its fundamental reliance on the non-specific 'seeking system' that acts as a 'gain' system and modulatory control system for all the prototype states, driving organisms out there to 'mix it up' with other living things, to find essential emotional and biological supplies. Major lesions of that core system consistently create akinetic mutism and a virtually total ablation of emotion (Watt & Pincus, 2004).

In any case, this simple typology for prototype emotion does not yield a large N (only seven major primary states and 'seeking system'). However, and quite suggestively, these break down into two obvious large clusters: an *Organism Defense Cluster — Fear, Rage and Disgust* (which protects us immunologically) vs. a *Social Connection Cluster* (*Lust, Play/Affection, Nurturance and Separation Distress*). This suggests the obvious conclusion (readily supported by any introspective review of human life) that outlining fundamental processes in the brain underpinning social connection and attachment are very critical territories for affective neuroscience. However, attachment is likely to represent a large umbrella of affiliated processes that bind creatures together, involving several mammalian prototype states (such as separation distress and rough-and-tumble play), and not one simple, uni-dimensional process. In any case, the joys of love and attachment, particularly the two classical and most intense manifestations of

attachment in romantic and maternal love, are among the very deepest gratifications that human beings strive to attain. Conversely, the catastrophic loss of a rewarding or deep primary attachment typically plunges us into some of the cruelest pains that human beings can ever experience. Such losses (including their symbolic variants, and cognized first cousins of separation distress found in intense guilt and shame) are far and away the most common precipitants for the induction of serious depressions.

Surprisingly in view of these considerations, the basic neural substrates of attachment, play, and social bonding are still relatively poorly charted in neuroscience, and certainly poorly charted relative to their central importance in human life. Attachment and its vicissitudes are still massively neglected in the models in psychiatry for disorders in DSM-IV. We probably know much more about visual experience and visual processing in the cognitive-sensory domain, or about fear in the emotional domain, than we do about attachment and its critical components. In a nutshell, neuroscience, outside of a few investigators such as Jaak Panksepp, has until recently (Bartels & Zeki, 2000; Bartels & Zeki, 2004) been reluctant to tackle the neural nature of love. This neglect may reflect effects of lingering behaviouristic assumptions on neuroscience yet to be fully mitigated. Indeed, the more personal, intimate and emotionally profound the experience, the less neuroscience really knows about it. Related to this has been a reluctance to consider that attachment and love are hardly uniquely human qualities (although their higher cortico-cognitive resonances likely are), and that their phylogenesis must go deep into our genetic antiquity. One of the most critical and least appreciated aspects within this large domain of social connection processes has to be the problem of empathy, one of our most critical social abilities, and essential to the mitigation of human suffering.

The Problem of Defining Empathy: Blind Men Inspecting the Elephant

The term empathy was originally translated by Titchener in 1909 from the German term '*Einfühlung*', a term from aesthetics meaning 'projecting yourself into what you observe.' There has been endless hairsplitting in various psychological literatures about the possible distinctions one might draw between the terms compassion, sympathy, and empathy. It is this author's contention that for the most part these finely nuanced distinctions offer little to empirical neuroscience, and additionally, that all terms outline a common ground of a *positively valenced supportive response to the distress of another creature*, and that defining this core process is the critical issue for neuroscience. Unfortunately, extensive literatures on empathy offer significantly different and even conflicting definitions, with conflicts typically centring on more cognitive vs. more affective emphases. One cluster of literatures emphasizes empathy as dependent upon perception of affective states, theory of mind, conscious imitation, perspective taking and the like, the other group emphasizes the centrality of affective activation in the empathizing subject. It has also been approached as an essential cross-species mammalian ability, as a more restricted higher cognitive function dependent on

theory of mind, and as a conditioned/learned social behaviour. These differential concepts of empathy often lead to disparate methodologies and confusing empirical results (see Decety & Jackson, 2004; Preston & deWaal, 2002 for reviews). Empathy challenges us to develop models from the standpoint of its development in both phylogeny and ontogeny, in terms of its basic adaptive value(s), and in defining its fundamental neural mechanisms. Although much literature presents a more cognitive view of empathy, the etymology of the word suggests that the term empathy be reserved for phenomenon that extend beyond a cognitive perception of another's emotional/internal state (although those may be 'cognitive components' of empathy and critical social cognitions). Empathy, and its closest synonym 'compassion', both denote that one is 'suffering with another'. Additionally, although this is somewhat more controversial, I would argue that empathy in this core affective sense must also contain an intrinsic motivation to mitigate the distress of the other party. As Decety & Jackson (2004) emphasize, many if not most definitions of empathy involve three central components: (1) feeling what someone is feeling; (2) knowing what someone is feeling; (3) having some intent to mitigate their suffering. Although this review will emphasize a more basic affective definition of empathy, empathy shares a fundamental border with emotion identification, and with emergent higher cognitive processes involved in perspective taking, theory of mind, and many aspects of social cognition.

General Conceptual Frameworks and Organizing Hypotheses

When trying to understand processes happening within the staggering complexity of adult human brains, there is heuristic value to first developing an animal model of a target phenomenon, and testing this against the more complex process in humans. Obviously, there are many complex cognitive processes and behaviours for which this is not possible, such as the primary use of language, religious experience, meditative states, etc. However, a relatively simple animal paradigm with face validity provides an initial template for the study of empathy in humans: the response of maternal mammals to separation distress signals (and other kinds of distress and biological need states) in their infants. This animal model leads directly to several core hypotheses, although I will acknowledge that these are not universally conceded in behavioural and affective neuroscience:

(1) Based on the universal mammalian phenomenon of nurturance and maternal care, most mammals presumably have some primitive empathic capacities.
(2) Empathy appears phylogenetically coincident with the social signaling functions of emotion and with the deepening of formation of social bonds.
(3) Therefore, it seems reasonable to suspect that primitive empathic ability developed concomitant to strong attachment, that evolution in some sense carved the two processes jointly, with social bonding being

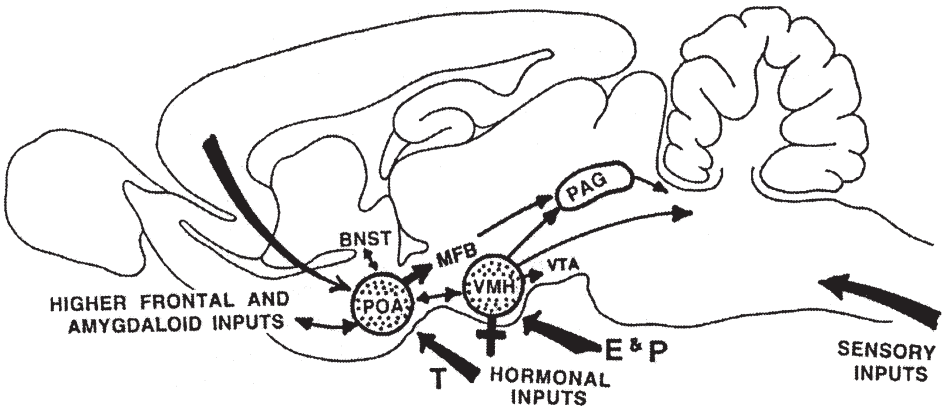

Figure 1a: Sexual Arousal Circuits

Lateral view of the rat brain summarizing two major areas that provide differential control over male and female sexual behaviours. Males contain a larger POA, and this area is essential for male sexual competence. The VMH is clearly more influential in female sexual responsivity. The systems operate, in part, by sensitizing various sensory input channels that promote copulatory reflexes. The extent to which these circuits control the affective components of sexual behaviour remains uncertain.

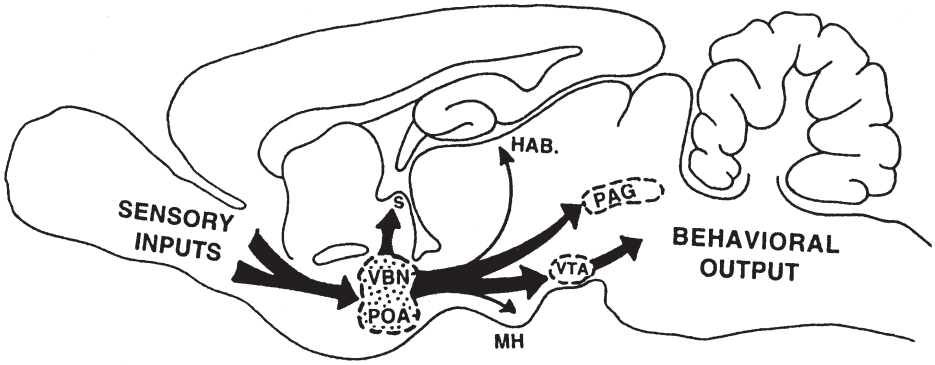

Figure 1b: Maternal Behaviour Circuits

General overview of maternal behaviour circuits in rodents. The central integrator is in the dorsal preoptic area (POA) and the ventral bed nucleus of the stria terminals (VBN), which receives various sensory cues for maternal behaviour and distributes controls into widespread brain areas, including the medial hypothalamus (MH), the ventral tegmental area (VTA), the periaqueductal gray (PAG), the habenula (HAB), and the septal area (S). The precise functions of these various areas remain to be identified.

Figure 1: Neural Systems for Sexuality and Nurturance (Panksepp, 1998)
(with permission from Oxford University Press)

critically enhanced by the ability to perceive, and relieve, the distress of a con-specific.

(4) The more complex phenomena of human empathy may reflect a 'cognized extension' of this mammalian prototype state of nurturing behaviour towards young, especially distressed young, but these critical cognitive extensions potentially allow for increasing appreciation of the internal spaces of others, and creation and extensive development of a theory of mind and perspective taking in human ontogenesis. However, these cognitive developments do not replace the largely subcortical prototype of mammalian nurturance, but rather reflect increasingly complex extensions of this process into telencephalic regions through distributed networks.

Animal Model Architectures: Distributed Networks for Social Connection

Consistent with the above noted fundamental division in the emotion taxonomy between *organism defence* and *social connection*, animal work (summarized in three graphics here) suggests major overlap in functional networks in most mammals subserving sexual arousal, separation distress/social bonding, and for nurturance and maternal care (see Panksepp, 1998 for extended summary). The animal data on maternal care and nurturance suggest an initial hypothesis that primitive empathic ability might be organized by basic systems subserving a complex of attachment-related processes. The networks implicated in the animal work on nurturance and attachment involve preoptic areas of the hypothalamus, ventral portions of the bed nucleus of stria terminalis, and ventral septum, with likely secondary roles played by other basal forebrain, diencephalic and midbrain systems such as habenula, and other hypothalamic areas. The preoptic and ventral bed nuclei appear to be the primary organizers of maternal behaviour. Major lesions to either of these regions devastates the capacity of female mammals to care for their young. This functional network may change in phylogenesis, with more involvement of paralimbic areas, consistent with Jacksonian principles. Evidence argues that in primates and hominids the anterior cingulate and nucleus accumbens are increasingly critical for attachment and separation distress and maternal behaviour. Evidence suggests that the various prototype emotional states all have basic architectures that funnel down into different portions of periaqueductal gray and that different prototype states may reflect differential columnar activity in PAG (Watt, 2000).

Other relevant experimental results might include findings that mu opioid receptor knock-out mice are deficient in attachment behaviours, with evidence that both approach and orienting behaviour towards mother and separation distress responses are attenuated (Moles, Kieffer, & D'Amato, 2004). This suggests that the neuropeptides regulating attachment and maternal care particularly oxytocin (Uvnas-Möberg, 1998; Carter, 1998), opioids, prolactin (Panksepp, 1998) have broad relevance for regulating empathic responsiveness, but there

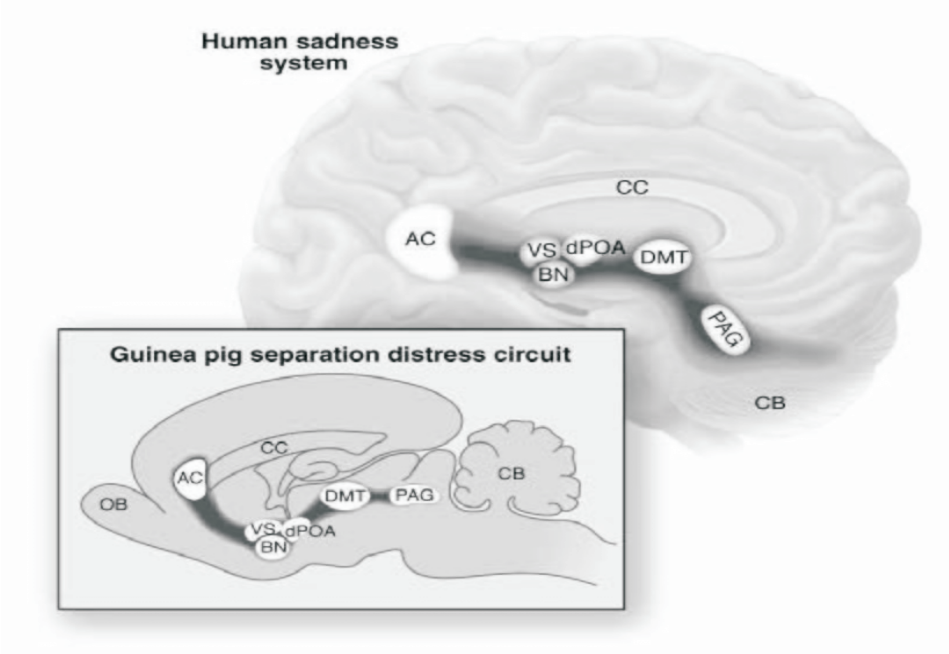

Figure 2: Human and Mammalian Separation Distress Systems
(reprinted from Panksepp, 2003 with permission of
American Association for the Advancement of Science)

has been to my knowledge no formal study of neuropeptides in human empathy. Jaak Panksepp (2004) notes three remaining major questions:

(1) Does the activation of distress circuits in young and relatively helpless animals evoke resonant activity in the same circuits of nearby adults?
(2) If such perceptually induced resonance does exist, is the evoked activity especially strong between strongly bonded individuals?
(3) Does such brain activity arouse caregiving in adults?

The default hypothesis here, suggested by the basic phenomenology of empathy, is that the questions above would be answered in the affirmative.

Empathy Functional Imaging Studies

There have been several studies of empathy or closely related phenomena using functional imaging. Unfortunately there have been different models of empathy, and different methods for testing and inducing it, making cross-study correlations and extrapolations difficult. Additionally, several basic caveats about functional imaging studies probably apply that cognitive neuroscience tends to minimize: (1) functional imaging (probably more true in fMRI than in PET) tends to minimally accent ventral brain regions (excepting cerebellum); (2) these

STUDY	METHODS	RESULTS/REGIONS OF INTEREST
Singer *et al*., 2004	fMRI study examining neural activations associated with witnessing pain in one's significant other vs. experiencing pain one's self. No pain in self and other is the control condition.	Empathy condition shows similar but not identical activations in various limbic structures as actual pain condition, but no sensory activations in somatosensory cortex (SII) seen in pain state. Empathy state shows activation in anterior cingulate, bilateral insula extending into left lateral prefrontal cortex, bilateral occipital cortex, fusiform cortex, and lateral cerebellum.
Carr *et al*. 2003	fMRI study modeling empathy as based in putative linkages between action representation (particularly imitative components) and emotion. Measured fMRI neural activations while subjects passively viewed and also imitated affective facial expressions, including with 'fractured' stimuli (just eyes, just mouth vs. whole face).	Right anterior insular and bilateral 'mirror' areas (area 44) inferior prefrontal system activations when subjects had to imitate or observe affective facial expressions, or even part-expressions (eyes vs. mouth), but imitation generated significantly greater peaks in ROI.
Lorderbaum *et al.* 2004	fMRI comparing maternal response to standard infant cry vs. response to mother's own infant crying.	Mother's own infant cry differentially activated cerebellum (esp. vermis), anterior and posterior cingulate, preoptic area/ventral bed nucleus, septum, bilateral insula, midbrain/ upper brainstem, right amygdala, and right temporal pole, hypothalamus, striatum, and thalamus.
Bartels & Zeki, 2004	Maternal love condition measured activations when mothers viewed their own child versus an age and familiarity matched acquainted child. Control for emotional valence: the same results were obtained when activity related to adult friendship was subtracted from maternal love.	Ventral/dorsal anterior cingulate cortex, frontal eye fields; fusiform cortex; bilateral insula; (ventral) lateral prefrontal cortex; occipital cortex; orbitofrontal cortex; thalamus; striatum (putamen, caudate nucleus, globus pallidus); periaqueductal gray; substantia nigra, hippocampus. Deactivations were noted in amygdala, posterior cingulate, mesial prefrontal/paracingulate, middle temporal cortex, occipitoparietal junction and temporal pole.

Meta-analysis Table for
Functional Imaging Studies of Empathy and Attachment

are correlative pictures with uncertain causal implications; (3) there is a widespread tendency in the literature to tweak p. values to yield very small hotspots, but it is very doubtful that brain really works this way, as opposed to widespread and highly distributed networks in which many structures may be variably activated or inhibited, some more so than others; (4) studies typically gloss over individual variations, often quite large. The table below summarizes the major studies and shows a broad functional overlap in many structures that we know from convergent sources are classically involved in emotional processing, especially the basal ganglia, various portions of cingulate and insula, along with various upper brainstem and basal forebrain and diencephalic regions.

Singer *et al*. (2004) found that empathy for pain in another correlates with increasing activity in the anterior portions of the insula and anterior cingulate cortex, although slightly different from regions involved in the experience of pain in one's self, and without the somatosensory components being present. The differential 'self vs. other' foci in paralimbic structures suggest that the degraded state of virtual pain in empathy isn't 'exactly' the same as actual pain experience, even after subtraction of sensory components. Of course, this study doesn't address empathy to primary emotions but the 'proto-qualia' of pain. Also notable is that this study looked at empathy activations *in a primary attachment*.

Carr *et al*. (2003) modeled empathy as based in putative linkages between action representation (particularly imitative components) and emotion. They measured with fMRI neural activations while subjects viewed affective facial expressions, including with stimuli that were 'fractured' (just eyes, just mouth vs. whole face), finding insular and inferior prefrontal system activations. One obvious question is whether the stimuli (static pictures of classic emotion faces) are evocative enough to generate real empathy, vs. just social cognition. Gallese (2003) has similarly argued that mirroring and imitation, dependent on inferior posterior prefrontal and parietal neurons, are substrates for empathy. However, it is not clear whether this truly explains empathy, and may conflate instrumental action matching (which requires mirror neurons) with the affective state induction/matching critical to empathy, which may not. Carr *et al*. also equate conscious, voluntary imitation of emotion with the emotion induction processes in empathy. These are possibly quite different, as empathy may use contagion processes, which may involve *unconscious mimicry* — see section on contagion. However, it is not established that primitive contagion processes require mirror neurons, as those mechanisms may be more primitive (Barsade, 2002), sitting 'under' systems involved in conscious imitation. All of this illustrates major continuing problems with conceptual models and definitions of empathy, esp. differing cognitive/voluntary vs. affective/involuntary emphases.

Lorderbaum *et al*. (2004) performed a functional MRI study examining maternal responses to a recording of the mother's own infant crying, comparing these responses to a recorded standard infant cry, with cerebellum (esp. vermis), anterior and posterior cingulate, preoptic area of hypothalamus, septum, midbrain, amygdala, striatum, and thalamus as regions of significant activation. Lorberbaum *et al* (2004) validates a more subcortical view of prototype emotion,

as systems critical to nurturance and maternal care in mammals remain centrally active in closely related states in humans. Although this was not a study explicitly about empathy, models reviewed here suggest major relevance for understanding empathy. This is further supported in the broad overlap of regional activations between the Singer *et al*. study and this study (basal ganglia, insula, cerebellum, anterior cingulate).

A study by Bartels & Zeki (2004) reveals very similar regions of activation when activations were measured in subjects looking at their child vs. a control image of another (known) child. The study also shows a conjoint transposition of results from an earlier study (Bartels & Zeki, 2000) on romantic love with the more recent study (Bartels & Zeki, 2004) on maternal love, revealing large overlap in functional networks (and also some differences) between the two most powerful forms of human attachment. Both activate BG, ventral striatum, cingulate, insula. These are all regions activated in the Lorberbaum *et al.* (2003) study of maternal responses to infant cries, *suggesting potentially close ties between attachment and basic empathic processes*, although conjoint mechanisms remain poorly outlined.

The table above suggests a *broad distributed network of structures and pathways essential to the activation of affective positive behavioural responses to loved and valued others, particularly those in distress, either partners or offspring*. The table outlines how regions activated by the distress of significant others show major overlap with regions implicated more globally in attachment and separation distress processes in both human imaging and animal lesion work, including particularly cingulate and insula, diencephalic and basal forebrain regions such as the bed nucleus of stria terminalis and the preoptic area, midbrain reticular activating system regions such as PAG, various striatal structures (probably ventral regions particularly) and the cerebellum (perhaps particularly midline portions).

Perception → Action Model (PAM) for Empathy from Cognitive Neuroscience

In the largest and most systematic review of empathy to date, Preston & de Waal (2002) reviewed an extraordinarily large body literature, and then presented a perception → action model (PAM) for empathy. Their model proposed that:

> the attended perception of the object's state automatically activates the subject's representations of that state; activation of these representations automatically primes or generates the associated autonomic and somatic responses. [ADD PAGE]

Although Preston and de Waal did a great service to the field in this ambitious review in bringing together into one source their coverage of an enormous body of work, along with many excellent contributed commentaries, several questions could be raised about this model. Challenges to this model might include:

(1) The model leaves out the intervening variable between perception and action of *emotion*.

(2) Does this model explain, or just restate, the basic explanatory challenge, namely that there is some kind of direct linkage between the receptive processing of, and activation of, an emotional state?
(3) What is the putative neurology and functional network(s) that might account for this linkage? Putative neural substrates were never clarified except for a 'diffuse perception-action mechanism'.
(4) The emphasis on 'representation' raises problems. Is emotional empathy dependent on cognitive 'representations' of an emotional state, or is it more a 'precognitive' process? Or potentially (and variably) both?
(5) Empathy clearly isn't automatic or 'reflexive' (Bandura, 2002), with many variables affecting its induction, so this model doesn't account for the multiple 'gating' aspects ('familiarity', cue salience, recent exposure to like trauma are findings all summarized in Preston and de Waal's review, but other variables may be intervening (see next two sections).
(6) What about empathy involving an intrinsic motivation towards reducing distress? This is totally omitted from the PAM.
(7) What about links to animal models and basic mammalian nurturance processes?
(8) What about the long intuited global relationship between empathy and attachment?

A Basic Model for Empathy: 'Gated Resonance Induction' of Another's Distress

Based in part on the above literature reviews, below is one possible model for major processing envelopes as necessary and sufficient conditions for the creation of empathy. The boxology model below emphasizes, first, receptive processing involved in appraisal and recognition of emotional states (these may be both cognitive and 'pre-cognitive'. Cognitive top-down components are presumably supplied largely by right hemisphere heteromodal systems and right somatosensory cortex (Adolphs *et al*., 2000; Adolphs *et al*., 2003) that process facial expressions, tones of voice, and body kinetics in terms of their affective content and meaning. There is evidence for emotion-specific contributions from various regions (insula/basal ganglia for disgust, amygdala for fear, while regionally specific contributions for recognition of sadness or separation distress and other prototype states have not yet been clearly defined). In the model, these heteromodal systems may feed into a poorly understood 'global gate' that controls relative activation of empathic states, with critical global gate variables determining the extent of 'resonance induction' for the subject, and the extent to which the 'variably degraded resonance state' is induced. This notion of resonance induction, unlike the PAM of Preston and deWaal, is not based on putative cognitive representationalist mechanisms. However, the model incorporates similar notions, in which receptive and expressive aspects of emotion are directly linked in a fashion still poorly understood. Additionally, 'resonance

induction' (the mechanism implicated in contagion) may reflect a fundamental property of all the primary emotional states, as *empathic inductions appear to have basic similarities with the mutual induction of playful affection and smiling responses* between two people, and even between two members of different species. Although processes by which playfulness, smiling, sexual arousal and nurturance/tenderness are mutually and reciprocally activated in virtually any human dyad remains largely a mystery in neural terms, recognition of basic 'resonance induction' goes back at least a hundred years in psychology, and in other literatures, perhaps much longer, even thousands of years. One might note a particularly prescient intuition by McDougall in his *Introduction to Social Psychology:*

> (there is) . . a special adaptation of the receptive side of each of the principal instinctive dispositions that renders each instinct capable of being excited on the perception of the bodily expressions of the excitement of the same instinct in other persons (McDougall, 1908).

Although there isn't universal consensus on this, without some kind of affective induction and associated motivation to relieve distress, one has to assume something less than a full empathic response, only a more cognitive appreciation of the other's affective state. In human empathy, however, the motivation to relieve suffering can be expressed through an enormously large N of potential behaviors, running the gamut from direct physical rescue from life-threatening danger, to careful listening and reflecting upon the other person's emotional dynamics and history in psychotherapy, and many shades in between. In the turgid and muddy water of the real social world, many social situations presumably show highly variable admixtures of empathy, accurate cognitive perception of another's emotional state (with or without empathic induction), and distorted perception or misrepresentation of another's intentions, internal states and emotional dynamics (see section on disorders of empathy). Such admixtures surely challenge both our scientific models (which must at least initially simplify the complexity of nature), as well as our own adaptive emotional capacities, which must differentially sort out these variably helpful and hurtful responses.[1]

Critical Variables Governing Global Gate For Activation of Empathic Responses

There are probably several if not many still poorly mapped variables that determine the intensity of an empathic response to the suffering of another. One might describe four fundamental classes of variables, some presumably genotypic (some version of 'native empathic ability'), some more phenotypic (*developed* level of empathic ability), some related to fundamental properties of the

[1] A study on empathy begging to be done would involve measuring neural activations in subjects attempting to formulate reasons why someone in a particular social scene might feel sadness and seeing if the accuracy or inaccuracy of this reconstruction mapped onto differential neural systems. One would speculate that various heteromodal regions classically associated with TOM would be essential, but that contributions of paralimbic regions might also appear.

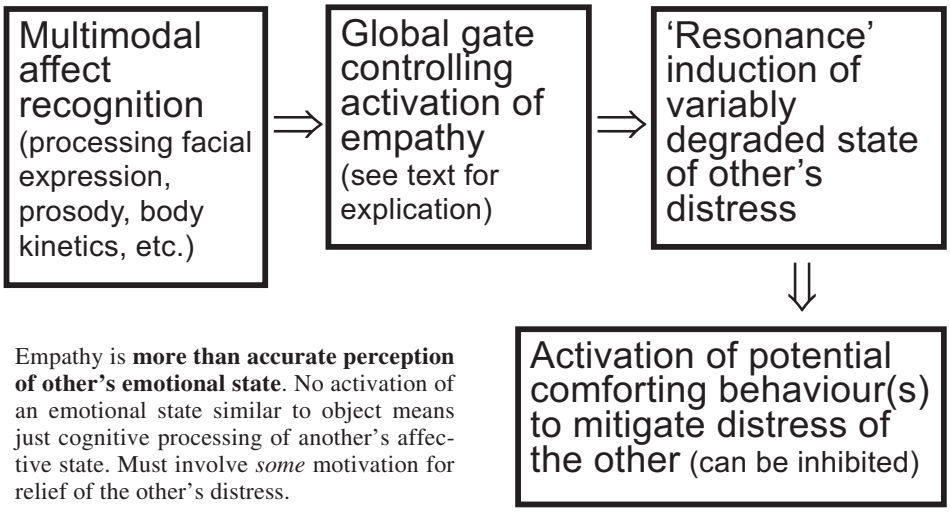

Figure 3: A Basic Model for Empathy

Many questions could be posed about each one of these boxes, their neural substrates, and any putative feedforward model such as this. For example, major questions can be raised about how much affect recognition is a conscious versus preconscious process, global gate variables for empathy induction are poorly articulated in the literature, and the links between resonance induction and behavioural activation towards comforting and relieving suffering remain a virtual mystery in neurological terms. Additionally, theory of mind processes may supplement or even replace direct perception of affect (see final section).

suffering party, and some 'state-dependent' (depending on the affective state of the empathizer). It seems very unlikely that the simple invariant/automatic PAM of Preston and de Waal can fully account for the enormous variability in our empathic responses to the manifest (and sometimes not so manifest) suffering of others; a more naturalistic model would have to model complex variables affecting empathic inductions. The model would have to centrally include both genotypic and phenotypic traits given the enormous variability in individual abilities to empathize. The animal literature well summarized by Preston and deWaal suggests that stimulus modality, familiarity with the distressed conspecific, similarity to the conspecific, and subject exposure to the same aversive event are variables, but this doesn't fully account for what we know affects empathic responses clinically and anecdotally. One might generalize from an enormous anecdotal database on empathy to hypothesize four primary classes of variables affecting empathic induction:

(1) First class of variables: native talent and developed ability. This is clearly very poorly mapped in the literature, but some evidence (Chartrand & Bargh, 1999) implicates some kind of *genomic variation in contagion/chameleon responses as a genotype* for empathy (also see Preston & de Waal, 2002). This poorly outlined genotype further

underlines the intrinsic tie of empathy to contagion (see later discussion). The phenotype would be the level of *developed ability for complex empathic responsiveness*. This presumably depends on interactions between the genotype and how empathically the person has been treated in their primary attachments, and is also poorly charted except anecdotally (Schore, 1994).

(2) Second class of variables: the degree of attachment to the object (subsumes the animal literature finding that 'familiarity' with suffering creature is an important variable). This emphasizes a likely 'proximate' mechanism for increasing empathic inductions. This is congruent with literally mountains of anecdotal data: the more we care for and love the suffering party, the more potent the resonance induction process is likely to be, *assuming that other variables don't contravene this,* such as our affective state or anger towards the other party. Imaging studies reviewed suggest that structures involved in primary attachments (Bartels & Zeki, 2000, Bartels & Zeki, 2004) have virtually total overlap with structures implicated in functional imaging studies of empathy (Singer *et al*., 2004; Carr *et al*., 2003) and maternal responses to infant cries (Lorderbaum *et al.*, 2004): ventral striatum, anterior cingulate, bed nucleus of stria terminalis, ventral striatum, preoptic area, and insula.

(3) Third class of variables: the degree of 'cuteness' of object, and the degree of felt potential vulnerability or helplessness of the object. This underlines the intrinsic tie of even more cognized human empathy to the mammalian prototype of nurturance: *empathy is fundamentally a protective and nurturing response to suffering in another, particularly those thought to be in need of care, and thought unable to fend for themselves*. We are far more empathically mobilized by the suffering of a small helpless creature with a rounded face and big eyes than by the suffering of a dominant, aggressive and powerful alpha male.

(4) Fourth class of variables: affective state of empathizer. In simplest terms, anger and emotional hurt towards another party tend to maximally inhibit empathy, along with many negative emotional states, particularly if they are intense. Indeed, rage transiently turns off empathic responsiveness almost completely. This reflects the basic 'gating logic' of prototype emotional states in which strongly negative states constrain activation of positives and visa versa, and empathy must come from a basically positive emotional stance towards the other party. How this 'valence gating' and reciprocal inhibition of prototype states is instantiated in the brain is unmapped (see Watt, 2000 and Bandler & Shipley, 1994 for one set of putative mechanisms). A common variation on this gating/ inhibition theme may be any version of moral judgment of the other's behaviour, with this representing a cognized expression of anger. Moral judgment is remarkably inhibiting of empathy, a point appreciated for many decades in the psychotherapy literatures. Last, but

not least, metabolic status (virtually ignored in the empathy literature) probably affects empathic induction. Simply put, we are typically lousy at empathy when tired and sleep deprived, hungry, or in pain, a point certainly not lost on virtually any overwhelmed parent!

In any case, the supposition of a 'global gate' controlling the activation of empathic responses has other predictive advantages, addressing one of Bandura's (2002) (and other commentators) chief concerns about Preston and de Waal's 'perception-action model' of empathy. It is also consistent with the multidimensional nature of virtually all affective activation, and with emotion being intrinsically 'valence-gated' (positives decreases activation thresholds for other positives and increase thresholds for negatives, and visa versa). Lastly, the model explains why the resonance state in empathy is 'variably degraded;' in situations where all of the global gate variables are largely maximized, we can experience a relatively more intense induction of the other's distress, a virtual copy as it were. Conversely, where the global gate is significantly constricted by any of those variables, one can easily predict a lesser to minimal or even absent degree of empathic activation. Additionally, it is worth noting that many of these variables also imply hidden variables/abilities in the empathizer, including reasonable affective regulation, self-non-self boundary (Decety & Jackson, 2004) and the importance of a positive emotional stance towards the suffering party.

Are Empathy and Attachment Interdigitating?

Much anecdotal clinical evidence from psychotherapy literature (Schore, 1994 for summary) suggests there is a likely massive reciprocity of attachment and empathy, and functional imaging studies just reviewed support this assumption, in terms of large overlap in activated systems and networks in empathy and attachment, although candidate shared mechanisms are still poorly plotted. However, there may be several basic functional interactions: (1) limit setting and modulation of behaviour: we can't afford to hurt the other too badly, or we lose them (they may detach and leave us); (2) basic developmental aspects linking phenotypes and genotypes of empathy: native empathic abilities are considerably strengthened (or partially eroded) by positive or negative early attachment experiences, and this long held belief in the psychological and psychotherapy literature has some support in the animal literature; (3) empathic responses when we are distressed enhance social bonding, while their absence makes attachments traumatic, chaotic and often short-lived (Schore, 1994). Empathy can also be fundamentally contrasted with common forms of 'distorted affective perception' that are basic concepts within psychodynamic theory indexing characterological pathology (e.g., concepts of 'transference', and 'projection'). These and related processes may be primary engines driving intermittent empathic failure and based in a kind of global analogical comparison between past and present, organized more in the right hemisphere (Watt, 1990).

Empathy in Psychotherapy and Psychoanalysis

The psychoanalysis and psychotherapy literatures are replete with literally many hundreds of discussions about the critical importance of empathy between therapist and patient, and the developmental literature in psychoanalysis (see Blanck & Blanck, 1979 for summary) contains literally many hundreds of references to the seriously damaging effects of early empathic failures in childhood attachments. Although reasons of space do not permit extended discussion of this vast collection of literatures, it is worth noting that empathy has been long hypothesized as a critical, *and possibly the most critical*, outcome variable from therapist side in the therapeutic interaction, in many schools of psychotherapy. The vast psychotherapy literature also emphasizes a more 'cognitive empathy', in terms of the therapist's intuition about cognitive and situational drivers for emotion, suggesting that complex cognitive operations (theory of mind, perspective taking, and emotion identification) are mixed in with more affective/empathic bridges to patients' internal distress.

Exploratory psychotherapies from the psychoanalytic tradition that centrally focus on the complex emotional currents and vicissitudes in the therapy relationship itself emphasize the pivotal importance of *empathic responses to and management of fundamentally non-empathic states.* What psychoanalysis has termed 'transference' (distorted perception of another's intentions/states based on similarity to early traumatic experiences) and projection (mis-attribution of one's own internal states to another) generate 'intermittent' empathic failures (see Langs, 1976 for overview). An empathic exchange around emotionally distorting perceptions/actions is seen as a central change agent in many psychotherapy metapsychologies, with this empathic management allowing the patient to reclaim and re-work traumatic experiences, and modify maladaptive defences and associated interpersonal distortions. In a nutshell, a large group of exploratory psychotherapies including psychoanalysis emphasize the *empathic repair of empathic failures* (Paivio & Laurent 2001). Perhaps no other literature, excepting perhaps the child development literature, underlines more categorically the centrality of empathy for creating positive personality and interpersonal dynamics than the wide ranging psychotherapy literatures. There is some evidence that mutually enhancing and reciprocal empathic failures in a dyad (what psychoanalysis termed 'transference — countertransference cycles') are generative for what Freud called the 'uncanny' ability of people to repeat the same traumatic scenarios over and over (Blanck & Blanck, 1974). Although careful operational definition of terms and empirical testing of constructs has typically not been a significant component of these literatures, such an emphasis should not go unnoticed.

Clinical Disorders of Empathy

Certainly any theory of empathy would have to account for its disorders, which can be broken down into several large clinical classes.

(1) Autistic disorders reflect a spectrum of failures (ranging from mild to catastrophic) in attachment/ empathic development, and suggests that capacities for attachment and empathy are fundamentally interdigitating in development. One suspects that theory of mind (TOM) may fail to develop secondary to failure of more basic and antecedent processes in empathy. However, the TOM literature has generally failed to consider primitive empathy/contagion as potentially precedent for TOM development.
(2) Effects of developmentally early (Anderson *et al*., 1999) and late orbital frontal lesions (Shamay-Tsoory *et al*., 2003) as well as the syndrome of sociopathy underline a crucial role for orbital frontal (OF) regions in empathy, although mechanisms for empathy deficit in orbital frontal lesion have not been clarified, and previous models emphasizing the OF role in internalization of interpersonal/social rules (Damasio, 1994) don't fully explain empathic deficits. Additionally, an implied component of the above model is intact affective regulation in the empathizing party, and this appears also dependent on orbital frontal function (Damasio, 1994).
(3) Massive right hemisphere insults (typically R MCA infarcts) tend to seriously blunt empathic ability (Shamay-Tsoory *et al*., 2003) , suggesting a primary role in empathy for the right hemisphere, particularly somatosensory systems, consistent with the work of Adolphs *et al*. (2000).
(4) Milder 'intermittent' empathic failures reside in all characterological disorders, beyond sociopathy. All characterological pathologies may 'systematize' certain characteristic forms of empathic failure, depending on the particular characterological disorder. More broadly, one suspects that all maladaptive defensive operations may potentially impair cognitive and/or affective aspects of empathy.

Contagion as a Core Component Process in Empathy: A Primitive 'Precognitive' Mechanism for Emotion Induction?

Although Preston & deWaal (2002) note contagion 'on a border with empathy', an emotional empathy depends on resonance inductions, suggesting that contagion is paradigmatic. Clearly, empathy requires more than simple contagion, in that the induced state of suffering must not flood the subject, but instead has to mobilize a helpful or comforting response, suggesting that affective regulation must be relatively intact in the empathizer, as must be self-other differentiation (issues noted by Decety & Jackson, 2004). Additionally, there must be a fundamentally positive affective stance towards the sufferer, in terms of the intrinsic motivation to relieve their suffering in some fashion. Thus, empathy cannot be *reduced* to contagion, but basic contagion may function as one of its core component pieces.

Contagion is perhaps most classically reflected in fear inductions in herd behaviour, but all of the prototype emotions appear to be 'catchy,' as playful, smiling, lustful, and tender responses all facilitate and activate the same states in others in close proximity, as of course do the prototype negative emotions of fear and rage (Hatfield *et al.* , 1994). Unraveling emotional contagion would give us critical insights into how different prototype emotional states generate affective resonances in others — resonances possibly enhanced by higher cognitive abilities, but that may not depend on them entirely. We sometimes don't even have to hear the joke the other party heard to start laughing ourselves, if the other's laughter is intense or 'infectious' enough — indeed even the language we use here speaks to the centrality of the phenomenon.

Both empirical investigations and theoretical overviews of contagion (Barsade, 2002) have also suggested that it may be an critical component of empathy. Sonnby-Borgstrom (2002) compared facial mimicry reactions, as represented by EMG activity when subjects were exposed to pictures of angry or happy faces, and the degree of correspondence between facial EMG reactions and their own reported feelings. Subjects in the high-empathy group were found to have a higher degree of mimicking behaviour, while those in the low-empathy group showed inverse zygomaticus muscle reactions, 'smiling' when exposed to angry faces. Arguing for a more primitive contagion component to empathy, the author concluded that differences between the groups in empathy appeared related to differences in automatic somatic reactions to facial stimuli, rather than to differences in conscious interpretation of the emotional situation. There is evidence that primitive contagion effects extend to synchronicity of autonomic states between empathizing subject and distressed object (Levenson, 1996), as a possible 'physiological substrate' for empathy. Play and smiling responses critical to early attachment of infant and mother (Bowlby, 1977; Trevarthen & Aitken, 2001) also appear mutually inducted via contagion mechanisms.

Interestingly, the emotional contagion literature suggests two different mechanisms, paralleling the cognitive vs. affective concepts about empathy: 1) a more subconscious, automatic, 'primitive emotional contagion' (Hatfield *et al.*, 1992); 2) more conscious and cognitive process (see Gump & Kulik, 1997), with this more 'cognitive contagion' typically reflecting more conscious imitation. Most work however suggests that contagion is typically defined by the more automatic, primitive processes. This primitive contagion occurs through very fast processes based in *automatic and continuous nonverbal mimicry and feedback* (Hatfield *et al.*, 1992, 1993, 1994), including *automatic, nonconscious mimicry of the other's facial expressions* (Lundqvist & Dimberg, 1995; Dimberg, 1982), *vocal tones* (Hatfield *et al*., 1994) and even *body language* (Chartrand & Bargh, 1999). These effects can be measured even for subliminal facial presentations (Dimberg *et al*., 2000). These effects appear to be *transmodal* (e.g, presentation of vocal affect and not just facial affect changes facial musculatures (Hietanen *et al.*, 1998). These unconscious mimicry effects have been found even in infants as young as a few days old (Field *et al*., 1982; Haviland & Lelwica, 1987). Indeed, one might wonder why such a central

process as emotional contagion has received so little systematic attention in affective neuroscience, and one is left to ponder the effects of a possible cognitive bias (that all emotion induction is top-down and fundamentally cognitive) and/or a too 'atomistic'- individualistic image of emotion (minimizing that emotion primarily structures prototypical interactions between creatures).

The implicit assumption that conscious 'slow' imitation must be using the same neural pathways as 'fast' unconscious imitation is widespread in the empathy literature, (see method of Carr *et al*., 2003), and contributes to the heavy emphasis on mirror neurons as paradigmatic for empathy (Gallese, 2003). However, the automatic and unconscious imitation underling contagion works on much faster time scales (Hatfield *et al.*, 1992; Dimberg *et al*., 2000) than the conscious imitation seen in mirror neuron studies. Indeed, there may be widespread conflation of the mechanisms of contagion with notions of both mirroring and 'shared representations' (that similar actions/states across both self and other are mapped to 'pooled representations' - Decety & Jackson, 2004). However, contagion developmentally precedes such shared representations, which don't start appearing until at least 18–24 months. A tempting speculation is that contagion mechanisms form a poorly understood 'developmental ground' out of which later arriving and more cognitive 'shared representations' or 'mirroring' phenomena develop, but there is regrettably little real data on this. Additionally, conscious action matching, while potentially linked to affective state matching, can't be synonymous with it, particularly if one respects distinctions between relatively voluntary instrumental action, and relatively involuntary affective action, another distinction rarely if ever observed in the empathy literature.

Additionally, although poorly plotted in terms of neural substrates, contagion reflects induced changes in activation of core structures for emotion. For example, masked fearful and angry expressions increase, and happy expressions decrease, amygdala activity (Morris *et al*., 1998; Whalen *et al*., 1998). It is tempting to speculate that the balance between primitive/automatic vs. more cognitively informed contagion shifts through ontogenesis, as evidence (Field *et al*., 1982) suggests that the more primitive and automatic contagion is gradually inhibited through ontogenesis, probably coincident with prefrontal system myelination and the development of increasing affective modulation/inhibition (Hsee *et al*.; Chemtob, 1992). However, some capacities for primitive contagion clearly remain even in adults. In this sense contagion may prime the system in a particular direction (say, fear), but based on later arriving cognitive appraisals, the system may not necessarily settle into a fear state, but may move in other directions. In this sense, the priming function of contagion is probably more powerful early in development and less so as time goes on, but still prepotent and probably underappreciated in cognitive and affective neuroscience.

Considerations from work by Adolphs (2002), suggest that these fast vs. long time scales reflect differential neural network activations, and that the short time scale activations recruit the superior colliculus and amygdala; both of them have extensive connections to PAG, and therefore (speculatively) are in a position to prime this core structure and our affective systems in a particular affective

direction. There is regrettably little explicit work on the neural substrates for primitive contagion, but Barsade (2002) and others have speculated that emotion induction comes from an afferent feedback process. As many facial, postural, and vocal feedback studies have shown, once people engage in mimicking affective behavior, they then often rapidly experience the same emotion associated with the original motor outputs being mimicked. Although a candidate mechanism for contagion remains elusive, it seems reasonable to assume that the classical action categories associated with primary emotions have the ability to excite more global resonances throughout the distributed architectures underpinning each of the prototype emotional states (but this remains poorly understood and in need of future study).

In any case, these considerations suggest that *affective vocalizations, facial expressions, and affective body kinetics jointly constitute a special privileged class of stimuli for the human brain*, and that the *brain's early (short time-scale) receptive functioning prioritizes contagion responses while later responses may prioritize more cognitive processing*. It also seems quite likely that there may be differential structures/ networks activated in fear contagion vs. play contagion vs. separation distress/sadness contagion, but much of this remains to be clarified and there is no empirical work that I am aware of looking at these issues.

Remaining Questions, Hypotheses and Potential Tests

A most basic question would be 'does the brain generate the primary receptive processing categories for the basic emotions out of the same neural cloth as the primary affective behaviours, and only later, through cortical ontogenesis, develop the fine grain of receptive ability that allows us to see these primary emotional categories in all their subtlety and range?' Recent evidence shows that *both recognition of fear in others and fear conditioning require the amygdala* (Adolphs *et al*., 2001), underlining intrinsic linkages between receptive processing/recognition, and conditioning to and activation of prototype emotional states. A key question becomes: *does a 'categorical' and 'preconscious'* (pre-cognitive?) *recognition of the prototype states of separation distress, anger, fear, play, etc., emerge from the same basal forebrain, diencephalic and paralimbic regions necessary for the core affective states?* If so, then cortical receptive processing and recognition of emotion (organized largely in the right hemisphere) may be a cortical-cognitive 'extension' of a more primitive receptive processing possibly invisibly embedded in the distributed subcortical emotion architectures. Although there is increasing interest on 'shared representations' (Decety & Jackson, 2004) and notions of a 'common neural code' linking action and perception as critical to empathy (as emphasized in Preston & deWaal's PAM), one has to wonder if the linkages revealed in contagion are potentially precedent for later developing cortical mirroring functions. Perhaps this primitive 'direct linkage' can be potentially accessed, even late in ontogenesis, in basic contagion phenomena, using mechanisms antecedent to and more primitive than traditional and better researched top-down

activation of emotion by appraisal and cognitive processing. Contagion work in human infants (Field *et al*., 1982) and with very short stimulus presentations (Dimberg *et al*., 2000), and animal work showing intact affective responsiveness in decorticate 'lower' mammals (Panksepp *et al*., 1994) all suggest the above hypotheses. This type of resonance induction may thus be a developmentally primitive induction mechanism that cognitive development largely but not totally supplants. This must be tempered with a critical caveat: Jacksonian developmental principles argue for increasing functional dependence of lower systems on dorsal cortical systems in both ontogenesis and phylogenesis. In other words, it remains to be seen whether even the most basic receptive processing of emotion in humans is, like movement, much more dependent on cortex than in other mammals.

Critical tests of proposed models would involve: (1) empirical testing and confirmation of the hypothesized four classes of variables controlling the 'global gate' for the activation of empathy; (2) clarification of relative contributions to emotional empathy from more primitive preconscious contagion type processes vs. conscious top down cognitive recognition of affective states, (3) clarification of differential networks underwriting fast, unconscious contagion (poorly mapped) versus more conscious imitation (better mapped and presumed dependent on mirror neuron areas [Gallese, 2003]); (4) clarification of mechanisms that link resonance induction of distress (whether from more cognitive vs. precognitive mechanisms) to helping/comforting behaviours (virtually totally unexplored).

Envelopes of Social Connection and the Challenge of the Cognition: Emotion Border

Empathy, play, and separation distress, have been mostly conceptualized in neuroscience as largely discrete processes. However, all of these may have been jointly selected as different threads woven into the full fabric of an increasingly social brain, different yet equally essential components of evolution's discovery of the major advantages in tying creatures together. Play and empathy both powerfully cement social bonds between affective creatures, and both may be critical to the regulation of mood. However, creating such a deeply social brain meant also increasing developmental vulnerability to loss and separation distress, deep emotional hurt, subsequent promotion of pathological defences, and increasing vulnerability to depression and low self esteem, given close links between depression and prolonged separation distress (Bowlby, 1977). In view of this, empathy may be nature's inoculation against these dangers, evolution's gift to balance the intrinsic deep emotional vulnerabilities of a highly social brain. Additionally, empathy enables the creation of 'intersubjectivity' and the increasingly social and shared nature of much of the content of human consciousness, in which individuals can have deeply shared emotional spaces, with this becoming a critical aspect of all long term attachments.

One is also left with the sense that like many phenomena on the cognition – emotion border, the concept of empathy is likely to be beset with controversies and confusion about its fundamental nature for some time to come. Indeed, the sprawling and disjointed literature on empathy resembles the blind men inspecting different portions of the elephant, defying the efforts of Preston and deWaal and many other reviewers to bring theoretical coherence to it. However, as they point out, the term is a global umbrella concept covering a large affiliated group of cognitive and affective processes and not a single unitary process. This suggests considerable complexities sitting on the vast emotion-cognition border, in terms of a continuum of affiliated and interactive processes: complex reciprocal influences between theory of mind constructions, emotion identifications, and affective resonance inductions (see graphic below). This continuum of processes on the emotion-cognition border is consistent with much evidence that the brain instantiates parallel mechanisms towards the same purposes, cutting across more ventral subcortical and more dorsal cortical systems. The later arriving cortical mechanisms supplement and modulate, rather than simply replace, the more primitive subcortical mechanisms, consistent with the seminal formulations of John Hughlings Jackson. One readily suspects that these two developmentally linked trends are largely responsible for the divisive fracturing of the empathy literature along 'cognitive *vs* affective' lines. Instead, it seems very likely that what we call 'empathy' naturalistically involves variable admixtures of more cognitive versus more affective activations.

In view of the skill of cortically intact humans in suppressing emotional displays, and how affect display prohibitions also potentially make others' feeling states anything but clear, interaction between these two major components may be adaptively essential. For example, one can find that careful cognitive analyses of another person's behaviour might reveal evidence of hidden distress and hurt, masked by defensive (or offensive!) behaviours that previously had elicited irritation or resentment. However, such additional perspective taking and review might activate an empathic response to a previously unseen hurt or emotional need, hidden completely from the empathizer's prior awareness, and this empathic response now supplants a previously irritated and judgmental position. Thus, fuller development of a complex and affectively sophisticated theory of mind, melded with more primitive resonance induction mechanisms, may allow us to empathize with suffering where it is not particularly visible. Such linkages between high-level cognitive processes and more fundamental affective contagion/resonance induction mechanisms may supply adaptive skill sets absolutely essential to the most effective therapists and parents (and much appreciated in our valued significant others and close friends!). Perhaps our most effective empathic responses cannot reflect a 'cognitive versus an affective' process but rather a potential linking of several affiliated processes:

(1) cognitive modeling of predisposing factors in the other person's personality, context/circumstances jointly determining potential activations of painful emotion in the other (a cognitively mediated awareness of

primary emotional 'drivers' and precipitators). This ability is ultimately dependent on complex, high-level conceptual knowledge of human emotional dynamics and an '*affective* theory of mind';

(2) other cognitive abilities involved in differential identification of prototype emotions, potentially driving internal access to knowledge bases about what typically precipitates such emotions;

(3) resonance inductions that would allow us to feel some of the other person's pain (whether perceived or inferred);

(4) an associated desire to mitigate their suffering in some fashion.

How these various processes might become linked together is unclear, but there is considerable evidence in the literature reviewed earlier that several paralimbic regions are likely involved, including several regions of anterior cingulate, posterior and anterior insular, and orbital frontal systems. In any case, it is likely that such complex processes emerge from the concerted working of multiple heteromodal, paralimbic, and subcortical systems. One also strongly suspects that a complex and optimal developmental trajectory for the brain is required to truly maximally develop empathy, in which the individual is both treated empathically, and is also expected to respect other's feelings and respond supportively to upset and pain. More primitive 'resonance induction' process also may developmentally bootstrap more cognitive processes in theory of mind/perspective taking; one hypothesis emerging directly from the above considerations would be that the failure of such bootstrapping is a critical dynamic in the cascade of developmental failures in autism. On the other hand, successful developmental instantiation of such adaptive operational linkages between resonance induction, a complex, affectively attuned theory of mind, and nurturing behavior may constitute what is referred to in folk psychology as 'wisdom and emotional sensitivity'. One suspects that this kind of complex multidimensional empathy is often times our best adaptive response as therapists, parents, mentors and friends to other suffering human beings. Thus, the very formidable challenge for a developing neuroscience of empathy is to bring our models asymptotically closer to the real, humbling complexity of nature — in this case, our human nature.

Cognitive Processes	↔	**Affective Processes**	
Theory of Mind	Emotion Identification	Affective Empathy	Contagion
Perspective Taking	Cognitive Empathy	Attachment	

References

Adolphs, R. (2002), 'Recognizing emotion from facial expressions: Psychological and neurological mechanisms', *Behavioral and Cognitive Neuroscience Reviews* **1** (1), pp. 21–61.

Adolphs, R. (2003), 'Cognitive neuroscience of human social behavior', *Nature Reviews /Neuroscience*, **4** (177), pp. 165–78.

Adolphs, R., Damasio, H., Tranel, D., Cooper, G. & Damasio, A.R. (2000), 'A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping', *Journal of Neuroscience*, **20**, pp. 2683–90.

Adolphs, R., Tranel, D. & Damasio, A.R. (2003), 'Dissociable neural systems for recognizing emotions', *Brain and Cognition*, **53**, pp. 61–9.
Anderson, S.W., Bechara, A., Damasio, H., Tranel, D. & Damasio, A.R.(1999), 'Impairment of social and moral behavior related to early damage in human prefrontal cortex', *Nature Neurosci*., **2**, pp. 1032–7.
Bandura, A. (2002), 'Reflexive empathy: On predicting more than has ever been observed. Commentary, Preston & de Waal (2002) Empathy: its proximate and ultimate basis', *Behavioral and Brain Sciences*, **25** (1), pp. 24–5.
Bandler, R. & Shipley, M.T. (1994), 'Columnar organization in the midbrain periaqueductal gray: Modules for emotional expression?', *Trends Neurosci*., **17**, pp. 379–89.
Barsade, S.G. (2002), 'The ripple effect: Emotional contagion and its influence on group behavior', *Administrative Science Quarterly*, **47**, pp. 644–75.
Bartels, A. & Zeki, S. (2000), 'The neural basis of romantic love', *NeuroReport*, **17** (11), pp. 3829–34.
Bartels, A. & Zeki, S. (2004), 'The neural correlates of maternal and romantic love', *NeuroImage*, **21** (3), p. 1155–66.
Blanck, G. & Blanck, R. (1979), *Ego Psychology 2: Psychoanalytic Developmental Psychology* (New York: Columbia University Press).
Bowlby, J. (1977), 'The making and breaking of affectional bonds: Aetiology and psychopathology in the light of attachment theory', *Br J Psychiatry*, **130**, pp. 201–10.
Carr, L., Iacoboni, M, Dubeau, M-C., Mazziotta, J.C. & Lenzi, J.L. (2003), 'Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas', *PNAS,* **100** (9) pp. 5497–502.
Carter, C.S. (1998), 'Neuroendocrine perspectives on social attachment and love', *Psychoneuroendocrinology,* **8**, pp. 779–818.
Chartrand, T.L. & Bargh, J.A. (1999), 'The chameleon effect', *J. Pers. Soc. Psychol.*, **76**, pp. 893–910.
Damasio, A.R. (2003), *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain* (Orlando, FL: Harcourt).
Damasio. A. (1999), *The Feeling of What Happens: Body and emotion in the making of consciousness* (New York: Harcourt Brace & Co).
Decety, J. & Jackson, P.L. (2004), 'The functional architecture of human empathy', *Behavioral and Cognitive Neuroscience Reviews*, **3** (2)*, pp.* 71–100.
Dimberg, U., Thunberg, M. & Elmehed, K. (2000), 'Unconscious facial reactions to emotional facial expressions', *Psychological Science*, **11**, pp. 86–9.
Dimberg, U. (1982), 'Facial reactions to facial expressions', *Psychophysiology,* **26**, pp. 643–7.
Field, T.M., Woodson, R., Greenberg, R. & Cohen, D. (1982), 'Discrimination and imitation of facial expressions by neonates', *Science*, **218**, pp. 179–81.
Gallese, V. (2003), 'The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity', *Psychopathology*, **36** (4), pp. 171–80.
Gump, B.B. & Kulik, J.A. (1997), 'Stress, affiliation, and emotional contagion', *Journal of Personality and Social Psychology*, **72**, pp. 305–19.
Hatfield, E.J., Hsee, C.K., Costello, J. & Weisman, M.S. (1995), 'The impact of vocal feedback on emotional experience and expression', *Journal of Social Behavior and Personality*, **10**, pp. 293–312.
Hatfield, E., Cacioppo, J. & Rapson, R.L. (1992), 'Primitive emotional contagion', in *Review of Personality and Social Psychology: Emotion and Social Behavior*, ed. M.S. Clark, **14**, pp. 151–77 (Newbury Park, CA: Sage).
Hatfield, E., Cacioppo, J. & Rapson, R.L. (1993), 'Emotional contagion', *Current Directions in Psychological Science*, **2**, pp. 96–9.
Hatfield, E., Cacioppo, J. & Rapson, R.L. (1994), *Emotional Contagion* (New York: Cambridge University Press).
Haviland, J.M. & Lelwica, M. (1987), 'The induced affect response: 10-week-old infants' responses to three emotion expressions', *Developmental Psychology*, **23**, pp. 97–104.
Hietanen, J.K., Surakka, V. & Linnankoski, I. (1998), 'Facial electromyographic responses to vocal affect expressions', *Psychophysiology*, **35**, pp. 530–6.
Hsee, C.K., Hatfield, E. & Chemtob, C. (1992), 'Assessments of the emotional states of others: Conscious judgments versus emotional contagion', *Journal of Social and Clinical Psychology*, **11**, pp. 119–28.

Langs, Robert (1976), *The Therapeutic Interaction: A Critical Overview and Synthesis*, volume two (New York: Jason Aronson Press).
Levenson, R.W. (1996), 'Biological substrates of empathy and facial modulation of emotion', *Motivation and Emotion*, 20, pp. 185–204.
Lorderbaum, J., Kose, S., Dubno, J.R., Horwitz, A.R., Newman, J.D., Kose, R., Hamner, M.B., Bohning, D.E., Arana, G.W. & George, MS. (2004), 'A third fMRI study of healthy mothers hearing infant cries', Poster, Neuroscience.
Lundqvist, L.O., & Dimberg, U. (1995), 'Facial expressions are contagious', *Journal of Psychophysiology*, **9**, pp. 203–11.
McDougall, W. (1908), *An Introduction To Social Psychology* (Methuen).
Moles, A., Kieffer, B.L. & D'Amato, F.R. (2004), 'Deficit in attachment behavior in mice lacking the mu-opioid receptor gene', *Science*, **304** (5679), pp. 1888–9.
Morris, J.S., Friston, K.J., Buchel, C., Frith, C.D., Young, A.W., Calder, A.J. & Dolan, R.J. (1998), 'A neuromodulatory role for the human amygdala in processing emotional facial expressions', *Brain*, **121**, pp. 47–57.
Paivio, S.C. & Laurent, C. (2001), 'Empathy and emotion regulation: Reprocessing memories of childhood abuse', *J Clin Psychol*., **57** (2), pp.213–26.
Panksepp, J. (1998), *Affective Neuroscience: The foundations of human and animal emotions* (New York: Oxford University Press).
Panksepp, J. (2003), 'Feeling the pain of social loss', *Science*, **302**, pp. 237–9.
Panksepp, J. (2004), 'Altruism and helping behaviors, neurobiology', in: *Encyclopedia of Neuroscience,* Third Edition, ed. G. Adelman & B.H. Smith (Elsevier).
Panksepp, J. & Watt, D. (2003), 'The ego is first and foremost a body ego', *Neuropsychoanalysis,* **5**, pp. 201–18.
Panksepp, J., Normansell, L.A., Cox, J.F. & Siviy, S. (1994), 'Effects of neonatal decortication on the social play of juvenile rats', *Physiology & Behavior*, **56**, pp. 429–43.
Preston, S.D. & de Waal, B.M. (2002), 'Empathy: Its ultimate and proximate bases', *Behavioral and Brain Sciences*, **25** (1), pp. 1–72.
Schore, A.N. (1994), *Affective Regulation and the Origins of the Self* (New York: Lawrence Erlbaum Associates).
Shamay-Tsoory, S.G., Tomer, R., Berger, B.D. & Aharon-Peretz, J. (2003), 'Characterization of empathy deficits following prefrontal brain damage: The role of the right ventromedial prefrontal cortex', *Journal of Cognitive Neuroscience*, **15** (3), pp. 324–37.
Shamay-Tsoory, S.G., Tomer, R., Goldsher, D. Berger, B.D. and Aharon-Peretz, J. (2004), 'Impairment in cognitive and affective empathy in patients with brain lesions: anatomical and cognitive correlates journal of clinical and experimental neuropsychology', 26 (8), pp. 1113–27.
Singer. T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, RJ. & Frith, C.D. (2004), 'Empathy for pain involves the affective but not sensory components of pain', *Science*, **303** (20), pp. 1157–61.
Sonnby-Borgstrom, M. (2002), 'The facial expression says more than words: Is emotional "contagion" via facial expression the first step toward empathy?', *Lakartidningen*, **99** (13), pp. 1438–42.
Titchener, E. (1909), *Experimental Psychology of the Thought Processes* (Macmillan).
Trevarthen, C. & Aitken, K.J. (2001), 'Infant intersubjectivity: research, theory, and clinical applications', *J Child Psychology Psychiatry*, **42** (1), pp. 3–48.
Uvnas-Möberg, K. (1998), 'Oxytocin may mediate the benefits of positive social interaction and emotions', *Psychoneuroendocrinology,* **23**, pp. 819–35.
Watt, D. F. (1998), 'Affective neuroscience and extended reticular thalamic activating system theories of consciousness', Target Article, *Association For The Scientific Study of Consciousness Electronic Seminar*.
Watt, D.F. & Pincus, D.I. (2004), 'Neural substrates of consciousness: Implications for clinical psychiatry', in *Textbook of Biological Psychiatry*, ed. J. Panksepp (New York, Wiley).
Watt, D.F. (2000), 'The centrencephalon and thalamocortical integration: neglected contributions of periaqueductal gray', *Emotion and Consciousness*, **1** (1), pp. 93–116.
Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B. & Jenike, M.A. (1998), 'Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge', *Journal of Neuroscience*, **18**, 411–18.