Bowlby's Attachment Control System Approach: An Alternative History of Emotion Modelling

Dean Petters [1] and **Everett Waters** [2]

Abstract.

Within the contemporary emotion modelling community historical reviews look back to an exchange between Ulric Neisser and Herbert Simon. Neisser (1963) claimed that the serial computers of the day could not effectively model emotion. Simon's (1967) response was that augmented with a variety of real time interrupt and goal-management mechanisms they could. This paper presents an alternative history of an emotional modelling approach developed by John Bowlby, in which he proposed emotional modelling mechanisms as a scientifically respectable updating of psychoanalytic theory. This approach places comparatively less emphasis on emotional interrupts and goal-management within a single locus of control, and respectively more emphasis on how multiple mixed and often conflicting goals can arise from conflicting internal working models which are formed at different stages of development. A key conclusion is the benefit for emotion modelling of the rich scenarios provided by attachment phenomena.

1 A FAMILIAR HISTORY

In 1963, Ulric Neisser [47] published a critique of psychological modelling. He recognised that computers provide valuable models of purposive cognition, such as reasoning. However, he challenged whether the modelling techniques of the day would provide adequate explanations of psychological phenomena like: the parallel operation of multiple motives; bored and distracted thinking; other kinds of minimally purpose cognition; unconscious processing; and above all, social decision making. Neisser viewed the cognitive modelling at this time as 'indomitably' single-minded and 'too purposive' to match the richness and diversity of cognition in humans. He pointed out that human information processing possesses a subtlety and flexibility. Neisser recognised the issue was partly to do with the distinction between serial and parallel processes. However, whilst he presupposed human cognition is parallel his criticism of the computational modelling of his day went beyond this assumption. He emphasised the depth as well as breadth of human cognition that provides spontaneity, multiple motives, mixed motives, and even boredom. In his view, the computers of that time did not possess the kinds of subtle, complex, often conflicted and contradictory motives that would distinguish them from one dimensional 'shallow fictional characters'.

Simon [64] responded to Neisser's challenge by accepting a general theory of thinking and problem solving must incorporate motives and emotions, and that the absence of such states was a major lacuna of theories at that time. He then went on to illustrate how serial processing can account for emotions as interrupts to ongoing serial processes by time-sharing of the processor between main tasks and interrupt mechanisms, allowing a processor to respond to urgent needs in real time. He also proposed a collection of goal-management mechanisms allowing a processor to manage goals through satisficing, setting aspiration levels, implementing impatience and discouragement functions, using queues and individual time allocations, and making choices among alternatives as responses to multiple criteria. These mechanisms go a long way to adding subtlety and complexity to the approaches which Neisser criticised. However, a key difference between Simon's and Neisser's approaches is, respectively, between a serial system that processes multiple objectives in a rationale manner through a single locus of control (and augmented with mechanisms to manage real time needs and multiple constraints), and a system with multiple loci of control and the kind of conflicting motives familiar to psychoanalytic theory. Neisser makes this point with explicit reference to the requirement for conflicted multiple motives operating at different levels:

> *"In the early days of psychoanalysis it was fashionable to devalue the obvious motives in favor of the unconscious ones, and to assume that cognitive activity was "nothing but" a way to placate instinctual demands. This tendency is happily no longer common: "rational" activities are unquestionably important in their own right to the person who engages in them. But we must be careful not to let the availability of computer models seduce us into the 19th century view of a man as a transparently single-minded and logical creature"*([47], p. 196)

Simon's approach to modelling emotions as interrupts to processing has been an inspiration for generations of AI researchers, with this approach being expanded and modified in various ways [69, 66, 37, 60, 63, 45, 44]. The current state of the field of computational emotion modelling has progressed a long way, but a limitation of the contemporary 'state of the art' is that it has still not developed a rich ecology of affective states like standards, values, preferences, attitudes, attachments, motives, moods, ambitions, and other various moral and aesthetic phenomena which are integrated with different kinds of emotional interrupts to provide the subtleties and complexities of humans that were highlighted by Neisser.

This paper is an attempt to provide a historical review of an approach to computational emotion modelling set out by John Bowlby in his development of the 'attachment control system' concept that originated as an attempt to explain psychoanalytic complexity. Bowlby's approach assimilated an armoury of modelling techniques from cybernetics, artificial intelligence and other scientific approaches to explaining phenomena of interest to psychoanalysis. This approach to emotion modelling did not originate with the ideas of Neisser or Simon (it would seem Bowlby was entirely unaware of

[1] Birmingham City University, UK, email: dean.petters@bcu.ac.uk
[2] SUNY, Stony Brook, USA.

this work), but instead developed directly from psychoanalytic theory and out of Bowlby's attempt to update this theory from the emerging sciences of the mind in a scientifically respectable manner. However, what Neisser, Simon, Bowlby and psychoanalytic theory all held to be true is that *"Human thinking begins in an intimate association with emotions and feelings which is never entirely lost."* ([64], p. 29).

2 ATTACHMENT WAS PRODUCED IN STAGES

Bowlby was spurred to develop Attachment Theory from his observation of the effects on human relationships of: war-time evacuation [5]; the prohibition of parental hospital visits to their young children [6]; the effect of early maternal deprivation on later development [7]; and the behavioural phases that are observed in long term separations and grief and mourning in infancy [12]. For approximately thirty years after the Second World War Bowlby's theoretical approach to explaining what we now term attachment phenomena progressed from initially explaining these behavioural patterns in psychoanalytic terms to invoking increasingly sophisticated information processing structures and mechanisms [19, 2, 23, 42].

During Attachment Theory's development Bowlby imported ideas from diverse neighbouring disciplines to explain this range of phenomena. Explanatory concepts which were newly prominent in the general academic environment had a particularly strong influence as Bowlby formulated elements of Attachment Theory at those times. As he elucidated details of Attachment Theory he adopted new concepts to substitute for particular psychoanalytic constructs which he wanted to transform. Surveying the range of concepts Bowlby introduced over time, we can make several generalizations. Firstly, he was influenced by the changes in the intellectual milieu between 1950 and 1980. His initial thinking in the late 1940s and early 1950s had a heavy psychoanalytic influence [10]; his earliest 'loans' were from Ethology [11, 15]; his latest from Cognitive Psychology; and in between he focused on concepts from Cybernetics and Artificial Intelligence [18, 20, 21, 19].

2.1 Before Attachment Theory: a psychoanalytic theory of attachment

Bowlby became interested in personality development and the key role played by an individual's early caregiving environment before he trained as a psychoanalyst ([2], p 333). His belief in the significance of real life events on the course of child development set him in conflict with several psychoanalysts with whom he worked at the Tavistock Clinic in the late 1940s, resulting in his forming his own research unit in 1948. At an early stage of his research career he chose to focus on the effects of early separation from the mother rather than other examples of disturbed family interaction. Bowlby made this decision due to practical considerations, as he intended to work within a scientific methodology which focused on analyzing the effect of environmental conditions. This was in contrast to adopting the retrospective case study conducted by other psychoanalysts. From Bowlby's view, separation events were particularly suitable for research as they were an event on record, whereas at that time there was no adequate reporting or documentation for other forms of disturbed family interaction ([2], p 334).

Bowlby's departure from the mainstream of psychoanalysis was also due to his awareness of problems with Freud's motivational theory. For Bowlby, this approach to motivation required revision because it was rooted in a drive theory which suggested infants were primarily focused on their inner drives and drive representations, and little interested in the social or physical environment per se. This focus inwards was in part driven by the psychoanalystic retrospective case study method which Bowlby had rejected. Critiques from psychology and philosophy of science also made clear that the drive theory of motivation was not tenable. It was not well supported by their own evidence, which itself was problematic, and seemed inaccessible to ordinary standards of empirical analysis and falsification [22].

In 1951 Bowlby published a landmark report to the World Health Organization which contained a substantial body of observations on the mental health of children [9]. Although this report was principally a survey of empirical work, it also illustrates the manner in which Bowlby was attempting to reshape psychoanalytic theory. Bretherton [23] notes that:

> *"it is interesting to examine the 1951 report from today's perspective. At that time Bowlby still used the terminology of traditional psychoanalysis (love object, libidinal ties, ego, and superego), but his ideas were little short of heretical."* ([23], p. 50)

An example of Bowlby's unorthodoxy which is highlighted by Bretherton ([23], p. 51) is where Bowlby explains that the creation of an infant's ego and superego are formed in a process of interaction with his mother:

> "[The mother] *is his ego and his superego. Gradually he learns these arts himself, and as he does, the skilled parent transfers the roles to him. This is a slow, subtle and continuous process, beginning when he first learns to walk and feed himself, and not ending completely until maturity is reached"* ([9], p. 53, quoted in [23], p. 51).

According to Bretherton, this description: *"sounds more Vygotskian than Freudian"*, ([23], p. 51). Why did Bowlby attempt a transformation of psychoanalytic terms like ego and superego rather than just rejecting all of the psychoanalytic framework outright? Psychoanalysis did possess a number of key insights into early experiences and relationships which Bowlby valued and wanted to maintain in his own approach [75]. He therefore wanted to reform psychoanalytic theory not replace it wholesale. However Bowlby, keenly aware of the sociology of science, recognized that critics would likely throw out the genuine insights along with the untenable motivation theory. He also recognized that the useful insights about early experience and relationships were logically independent of the drive theory. Central then to his work on attachment theory at this stage was to find an alternative approach to motivation. The key here was to avoid replacing one kind of magic (drives) with another (e.g., the infant intends, signals, wants, needs, loves, etc. the mother) thereby incurring what Dennett [29] calls "intelligence loans", presumptions of intelligence that are unlikely to be accounted for ("paid back") [61]. As we shall see, first in Ethology, and later in Cybernetics and Artificial Intelligence, Bowlby found approaches that could account for what he called 'the apparently purposeful' organization of observable attachment behaviour without incurring such intelligence debts. The example above, where Bowlby attempts to provide a causal origin for the ego and superego is merely an early example of what he later accomplished in a more radical and explicit manner.

As Ainsworth and Bowlby [2] recount, it was in the early 1950s that:

> *"Bowlby*[...] *had begun a search for adequate explanation of the empirical findings, having found none in current psychoanalytic theories to account for young children's responses to*

separation and reunion, or indeed how the tie to the mother develops. At this point Konrad Lorenz's work on imprinting became available in translation. Sensing its possible relevance to his problem and encouraged by Julian Huxley, Bowlby began delving into the ethological literature. [...] *During the early 1950s Bowlby was also deeply influenced by his membership in an international and interdisciplinary study group on the psychobiology of the child convened by the World Health Organisation, which met annually. Among the members were Piaget, Lorenz, and Margaret Mead, and among guest speakers were Julian Huxley, von Bertalannfy, and Erik Erikson."*([2], p. 337)

Bowlby's revision of psychoanalytic theory with an ethological perspective occurred gradually. The transitional nature of Bowlby's theoretical perspective at this time is illustrated by Bowlby himself in 1953, when he stated:

> *"I want to remark on three or four psychological processes which may be relevant, and in doing so I shall speak in a hybrid, bastard language which I have come to use, which derives from both psychoanalysis and ethology"* ([10], p. 183-184).

Bowlby did not experience an overnight conversion from a psychoanalytic view to the scientific view provided by the then nascent disciplines of cybernetics, artificial intelligence and systems theory. However, we can see from the meetings he attended and the interactions he engaged in at these meetings that he was inexorably moved towards viewing attachment relations in scientifically respectable information processing terms[3].

2.2 Attachment Theory proper: an evolutionary ethological theory

After Robert Hinde (a leading ethologist) joined Bowlby's seminar group at the Tavistock clinic in 1954 the incorporation of ethological theory in Bowlby's conceptual development deepened. This collaboration helped contribute to Bowlby's first formal statement of attachment theory in his 1958 paper: *'The Nature of the Child's Tie to his Mother'*[11]. The three main ethological concepts which Bowlby incorporated in this version of his theory were that: (1) attachment behaviour are a species-specific behaviour patterns which he termed 'instinctual responses'; (2) these behaviour patterns are activated and terminated by various external and internal stimuli and (3) these simple sequences of behaviour are integrated into more complex behavioural patterns ([11], p. 366). It is worth emphasising that although later theories provided more detail on how the attachment system develops, even this early theory presents attachment responses as not just preformed and waiting to be triggered or maturing without experience but rather constructed through interaction between infants and their caregiving environment[4]. The 1958 version of Attachment Theory, with ethological instinctive behaviour as a new substitute for the Freudian instinctive behaviour of Bowlby's previous approach, was in several respects only part-way towards Bowlby's final characterization of Attachment Theory.

2.2.1 Six papers between 1960 and 1963

Between his first presentation of Attachment Theory-proper in 1958, and the 1969 publication of the first volume of the Attachment and Loss Trilogy, Bowlby published six key papers that illustrate the movement of his thinking in this period. He published two papers concerned with empirical descriptions of separation anxiety [14, 16], and three papers concerned with mourning, particularly in infancy, [12, 13, 17]. In his 1960 paper: *'Ethology and the Development of Object Relations'*[15], Bowlby considered how ethological mechanisms can substitute for psychoanalytic concepts. For example, using ideas from ethology in understanding the ontogeny of object relations and anxious, depressed and defensive responses which are triggered by infant-mother separations. He considers the benefit of the ethology inspired idea that attachment is a primary need which leads to the emergence of reciprocal social relationships during ontogeny. This theoretical innovation leads to the expectation that distress and anxiety are normal expectations when separations occur, and that depression and separation anxiety can be seen as different stages in one process. Whilst he still refers to many psychoanalytic constructs like 'orality', 'repression' and 'symbolic substitutions' he does speculate whether behaviour previously thought to arise from these mechanisms might be better described by ethological concepts. He gives an example, where psychoanalysts would explain one activity taking the place of another because of symbolic equivalence whereas this may occur at an intrasymbolic level due to ethological displacement activities.

2.3 The final 'full-strength' control system version of Attachment Theory

2.3.1 Continuity and Change in theoretical developments

Bowlby reflected on the changing nature in his own theoretical approach during the 1960s, when he described the difference between his 1958 version of his theory and the version in the 1969 first volume of the Attachment and Loss Trilogy:

> *"The hypothesis proposed represents a development of that advanced by me in 1958. The principal change is due to better understanding of control theory and to recognition of the very sophisticated forms that behavioural systems controlling instinctive behaviour may take. In the present version of the hypothesis it is postulated that, at some stage in the development of the behavioural system responsible for attachment, proximity to mother becomes a set-goal. In the earlier version of the theory five patterns of behaviour - sucking, clinging, following, crying, and smiling - were described as contributing to attachment. In the new version these same five patterns are still held to be of great importance, but it is postulated that between the ages of about nine and eighteen months they usually become incorporated into far more sophisticated goal-corrected systems. These systems are so organised and activated that a child tends to be maintained in proximity to his mother.* [...] *The earlier version of the theory was described as a theory of component instinctual responses. The new version can be described as a control theory of attachment behaviour* ([19], page 180,)

[3] For example, the transcript of the second meeting of the 'World Health Organization Study Group on the Psychobiological Development of the Child', in 1954, records Bowlby and Grey Walter discussing the scientific nature of the psychoanalytic 'superego' construct:
BOWLBY: *The superego is rather complicated and contains more than one variable*
GREY WALTER: *Can you measure the superego?*
BOWLBY: *You cannot.*
GREY WALTER: *Then what is the point of discussing it?*
BOWLBY: *I think it is useful to try and see how things relate in these psychological functions after which we are dimly trying to grope.*
([38], p. 205)

[4] [73] provides a detailed account of Bowlby's introduction to, and adoption of, ethological theory.

Bowlby's newer 1969 version of Attachment Theory shows the continued importance of secure base behaviour with an increasing role for mental representation. As a control theory, the newer framework provides a greater focus on the attachment system as directed towards outcomes as set-goals to be achieved from a flexible behavioural repertoire rather than a system that simply involves triggering preset responses. However, the new theory still includes a strong ethological influence. Although the control systems formulation was a major departure from Bowlby's early instinct theory, he retained his commitment to behavioural biology. For example, Bowlby still presents the attachment system as an instinct to form bonds and as a system that is activated by species specific patterns of care. In addition, Bowlbys new terminology of behaviour systems only masks a core theoretical inheritance from his ethological instinct theory. As Hinde notes:

> *"The concept of a behavioural system is, in fact related to one meaning of the term instinct.* [...] *It has been used in a rather special sense by ethologists to refer to systems postulated as controlling a group of behaviour patterns that together serve to achieve a given biological end"*([41], p. 57).

The final version of Attachment Theory, which was set-out across all three volumes of Bowlbys Attachment Trilogy [18, 20, 21, 19], involves a richer and deeper conceptualization for attachment phenomena. Within a control system framework, Bowlby's three volume trilogy included information processing concepts such as homeostasis, hierarchical plans, internal working models, selective attention, and meta-cognition[5]. These concepts provide a supporting framework for explanations of a range of attachment phenomena. However, even whilst Bowlby was providing an alternative to psychoanalytic explanations for attachment behaviour, the range and type of phenomena of interest to psychoanalysts provided much of his research agenda:

> *"The resulting conceptual framework is designed to accommodate all those phenomena to which Freud called attention for example, love relations, separation, anxiety, mourning, defense, anger, guilt, depression, trauma, emotional detachment, sensitive periods in early life and so to offer an alternative of the traditional metapsychology of psychoanalysis"* ([19], page 668).

Although Bowlby claimed that the trilogy sets out an approach which had already been fully conceived of at its initiation, there were some limited changes in emphasis between the 1969 and 1980 publications. For example, the 1969 volume incorporated Cybernetic and AI concepts and the 1973 and 1980 volumes show a switch in emphasis to Cognitive Psychology. This change is not surprising. Bowlby noted that both his, and Freud's previous formulations of instinctive behaviour, were both *"a reflection of the scientific climate of the times"* ([19], page 18). Clearly, to incorporate up-to-date citations in the three volumes of his trilogy as he published them would involve inclusion of some influences from contemporary theoretical sources. However, how far Bowlby was influenced merely by the nature of the ideas in current circulation at any given time should not be over emphasized. During the 30 years in which he constructed Attachment Theory he selected particularly suitable concepts which matched his current requirements. So he did incorporate concepts from the current literature, but from a wide set of candidate ideas he selected those which possessed the appropriate properties and fitted within his existing framework. For example, in both the first and third volumes of the attachment trilogy psychoanalytic defensive processes are reframed in terms of how sensory inputs are processed. However, in the first volume this process is framed and referenced in terms of neurophysiological processes, in particular drawing upon the work of Magda Arnold ([19], page 103), whereas in the third volume the same attachment phenomena are described in the newly fashionable cognitive psychological terms of selective attention ([21], chapter 4).t

2.3.2 Control Systems

In the first volume of his Attachment Trilogy, Bowlby was continuing his search to consolidate the conceptual foundation for Attachment Theory by replacing Freud's concept of Psychical energy and its discharge ([19], p. 18). He introduced the control system concept for this purpose, and set out a range of control systems of increasing sophistication. For a simple example of a control system, which acted as a regulator by keeping a single variable constant, Bowlby presented the humble thermostat. He noted that this system was relatively static (with a pre-set goal) and unable to act as a model of even the simplest form of instinctive behaviour. However, Bowlby ([19], p. 44) showed how this control system design might be elaborated by allowing its goal setting to be determined by another control system. He also sketches out further extensions to this simple design of a greater sheer scale and complexity. In Bowlby's thinking the importance of the intimate integration of different representational forms we have set out is demonstrated when he noted:

> *"The mental apparatus can be thought of as made up of a very large number of complex control systems, organized in a loosely hierarchical way and with an enormous network of two-way communications between them. At the top we postulate one or more principal evaluators and controllers, closely linked to long term memory and comprising a very large number of evaluation (appraisal) scales ranged in some order of precedence. This system, or possibly federation of systems, I shall call the Principal System (s), this leaving open the question whether it is best regarded as singular or plural"* ([21], p. 52).

This passage is one of the more explicit descriptions by Bowlby for what contemporary researchers in cognitive modelling would term an architectural solution to the problem of understanding attachment phenomena. Though it is not a detailed description it does direct research on attachment modelling to the exploration of how different constituent parts for an attachment control system might be integrated. Newell (1990) defined a cognitive architecture as: *"the fixed (or slowly varying) structure that forms the framework for the immediate processes of cognitive performance and learning."* ([48], p. 12). So Bowlby's conception of the attachment control system is certainly a cognitive architecture in this sense. The temporal structure of attachment styles as long term control states, and related shorter term states in particular attachment related episodes, can be compared with other kinds of affective control states within cognitive architectures [67]. Anderson [3] provides an outlook on cognitive architectures that considers them formed of three factors: their function; the structures and mechanisms they possess; and the computations they perform. In the first volume of his Attachment Trilogy, Bowlby did not just match interesting behaviours with potential information processing structures and mechanisms. He also spent a great deal of this

[5] Van der Horst [73] described Bowlbys adoption of an ethological framework as an *"Archimedean moment"*[73], p. 3). However, the adoption of a cybernetic and control systems framework in the late 1960s has as much claim to be a core pivotal moment as the earlier adoption of ethology.

first volume concerned with explicating the evolutionary function of attachment behaviour. Bowlby set out all the three required ingredients for an architectural analysis of attachment phenomena, namely: empirically observed attachment behaviours; information processing structures and mechanisms; and an evolutionary functional analysis that matches these structures and mechanisms with the observed behavioural patterns. However, since Bell and Newell [4] didn't coin the term 'cognitive architecture' until 1971 it is understandable that it was not adopted by Bowlby when he originally formulated Attachment Theory, who instead used the term 'attachment control system'.

2.3.3 Action Selection

Bowlby did not discuss how particular actions or behaviours were selected using the contemporary term 'action selection'. However, he did present ethological mechanisms operating within the attachment control system architecture as carrying out this role. In both the 1958 and later versions of Attachment Theory, the behaviours related to attachment were organized according to four behaviour systems, the attachment, fear, sociability and exploration systems. According to Bowlby, what defines the attachment control system is not a set behaviour repertoire but the outcomes that predictably follow from these behaviours. Similar behaviours may be produced by different behaviour systems. In addition, each behaviour system has its own ontogenetic development, initially producing reflex actions and later in infancy producing fixed action patterns which increase in the complexity of their organisation in sequences and chains. Where in the 1958 theory attachment instincts were linked to the activation of particular behaviours (which can be viewed as having an implicit goal in the sense of having a predictable outcome in environments similar to the species environment of evolutionary adaptedness), in the later theory they were linked to the achievement of particular goals which are explicitly held as representations which can be achieved by a variety of actions.

2.3.4 Hierarchical Planning

In the first volume of his trilogy, Bowlby also set out differences between control systems in terms of how the behaviours within them are organized. He presented behavioural chaining as an example of a simple organizing principle for control systems, and hierarchical planning as much more complex and flexible ([18], p. 76). In this approach, plans are composed of sub-plans, and each plan and subplan is a set of instructions for action. So a high level plan can give a main objective and general strategy, where subplans deal more with the details of how to implement actions. However, Bowlby presents a very broad range of planning types. He includes planning examples of both explicit human plans and planning carried out by rats and other animals [18], p. 79-80). Although Bowlby presented contexts in which different kinds of plans would be formed, he did not distinguish these planning examples in terms of the sorts of representational or computational details needed to implement running simulations.

2.3.5 Internal Working Models

The concept of 'Internal Working Models' (IWMs) as important representations within the attachment control system was also introduced in the first volume of the attachment trilogy [18]. These are described by Bowlby as higher level representational forms which integrate and exert control over lower level control systems. Their principal information processing function is to allow predictions to be made about the likely outcomes of taking actions within a given environment, and Bowlby notes: *"the two working models each individual must have are referred to respectively as his environmental model and his organismic model* ([18], p. 82). These two models transmit, store and manipulate information and allow the individual to *"conduct small scale experiments within the head*" ([18], p. 81). Their function, in terms of Bowlby's agenda of reforming psychoanalytic theory, was to take the place of the internal worlds of traditional psychoanalytic theory. In this first volume of the trilogy Bowlby emphasizes the requirements for IWMs to be updated. He also briefly observes that pathological sequelae of separation and bereavement can be understood in terms of out of date models or half revised models which may contain consistencies and confusions ([18], p. 82). Bowlby invokes IWMs at early stages in development and later on, when linguistic skills and conscious reflection can enable models to become more adequate ([18], p. 84). As with his presentation of planning, he does not emphasize the representational details for IWMs.

Bowlby linked the construction of plans with the operation of Internal Working Models:

> *"Not infrequently, many alternative plans are concocted, their potential consequences imagined (on the basis of models of environment and organism) and the consequences of each plan appraised. Only after that is any particular plan put into operation"*([18], p. 114)

Petters [55] presents a historical review of the development of the IWM concept in Bowlby's writing and shows the change over time from an analog to symbolic conception of IWMs. More recently, artificial neural networks, have been used to simulate how attachment representations may change or remain stable over development. A number of researchers have used Hopfield Artificial Neural Networks to simulate prototype formation [70, 32]). Edalat and Mancinelli [32] present a model that explains early attachment stability as arising from strong patterns with large basins of attraction in Hopfield neural networks. Fraley [36] uses a synchronous auto-associator network to investigate how early relationship prototypes respond to differing schedules of subsequent experience. A key conclusion from these artificial neural network simulations is that early prototypes are not over-written, and so show greater continuity, when new relationship experiences are inconsistent. But consistent presentation of new prototypes does result in gradual change.

2.3.6 Stability and lability over ontogenetic development

Bowlby didn't use the terms 'precocial' and 'altricial' to describe stable or labile control systems, but the developmental trade-offs he sketched for these contrasting systems match current conceptions in research on altricial and precocial patterns of development in contemporary computational cognitive modelling [26]. For example, Bowlby [18], p. 46) presents a developmental trade-off whereby ontogenetic lability in a control system might result in a longer developmental duration, but may also result in this control system becoming better adapted and more flexible than a stable fixed alternative. Recent work has framed similar ideas in terms of altricial and precocial forms of development, within artificial systems which might be used in computational modelling [26]. However, it is notable that this recent work has additionally linked altricial development (which corresponds to Bowlbys labile developmental pattern for control systems) with the development of higher level more explicit representational forms.

2.3.7 *Representational change through ontogenetic development*

Bowlby compared changes in the kinds of representations which infants and children would possess in their attachment control systems with representational forms such as sensorimotor, pre-conceptual and symbolic representations which were theorized to occur in different Piagetian stages ([18], p. 153). He also speculated that actions which are automatically triggered early in development, such as feeding reflexes, might then be organized into causal hierarchies which can then form part of plans with set-goals. ([18], p.160; [57], p. 54-55). This appreciation of how representational forms can change early in infancy was matched by an acknowledgement of the important role played by language and symbolic forms of representation in later stages of development:

> *"Thus whereas during infancy and childhood humans are incapable of structuring their behaviour in any way more complex than the simplest of plans, in adolescence and adulthood behaviour is habitually structured on the basis of elaborate plan hierarchies. This tremendous development on the sophistication of the behavioural organizations employed is made possible, of course, by the increasing capacity of the growing human child to use symbols, especially language."*([18], p. 155)

Bowlby recognized that natural language is the ultimate and most sophisticated way in which an individual can represent themself within their social environment. This form of representation has the benefit that *"instead of each one of us having to build his environmental and organismic models entirely for himself, he can draw on models built by others"*([18], p. 82)[6]. A benefit of non-communicative aspect of language is that the possession of language allows more flexible and imaginative plans and sub-plans to be created, and shared with others, constituting a possible form of therapy.

Significantly, Bowlby also recognized that the nature of early representations are intimately linked with later patterns of behaviour, for example:

> *"Man's capacity to use language and other symbols, his capacities to plan and build models, his capacities for long-lasting collaboration with others and for interminable strife, these make man what he is. All these processes have their origin during the first three years of life, and all, moreover, are from their earliest days enlisted in the organisation of the attachment behaviour."*([18], p. 358)

2.3.8 *Homeostatic control, purpose and teleology*

In the second volume of the Attachment Trilogy, Bowlby adopted the biological concept of homeostasis and applied it to behavioural as well as physiological control systems. In this view, physiological homeostasis which regulates food and sleep are an inner ring of control. The attachment system constitutes an outer 'behavioural' ring which is a complement of this inner 'physiological' control system (Bowlby 1973, chapter 9). Bowlby discussed how Sommerhoff's ideas in *'Analytic Biology* [71] provided an explanatory framework which showed how apparently goal-directed behaviour can result from physical systems without the presence of rational agents, conscious mental processes, or explicitly held goals ([71], p. 66). Petters [55] compares the approach taken by Bowlby to homeostatical control by the attachment control system to the autopoetic approach in contemporary enactivism.

2.3.9 *Defensive processes, consciousness and model construction*

Bowlby explained defensive processes in terms of multiple working models. Some working models are open to conscious access. Some are unconscious but still highly influential, and so act as *"a version in different terms, of Freud's hypothesis of a dynamic unconscious"* ([20] p. 238). Bowlby also suggests IWMs can be formed with incompatible information, some of which becomes dominant and exerts influence either consciously or unconsciously:

> *"In a person suffering from emotional disturbance it is common to find that the model that has the greatest influence is one that developed during his early years and is constructed on fairly primitive lines, but that the person himself may be relatively, or completely, unaware of; while simultaneously there is operating in him a second, and perhaps radically incompatible, model, that developed later, that is much more sophisticated, that the person is more nearly aware of and that he may mistakenly suppose to be dominant."* ([20] p. 238).

In this passage Bowlby is setting out forms of interaction and differences in representation which provide another strong requirement for contemporary models of attachment. Bowlby also linked the operation of information processing with conscious awareness:

> *"Reflection suggests that many of the mental processes of which we are most keenly conscious are processes concerned with the building of models, with revising or extending them, checking them for internal consistency, or drawing on them for making a novel plan to reach a set-goal. Although it is certainly not necessary for all such processes always to be conscious, it is probably necessary that some should be so sometimes. In particular, it seems likely that revising, extending and checking of models are ill done or done not at all unless a model is subjected from time to time to whatever special benefits accrue from becoming conscious"*([18], p. 82).

The expectation that IWMs are sometimes open to conscious reflection provides a further strong constraint for computational models.

2.3.10 *Therapy as meta-management*

In the third volume of the Attachment Trilogy [21], Bowlby drew for new inspiration upon material from the emerging field of Cognitive Psychology. He explained Freudian defensive processes in terms of selective attention ([21], chapter 4), and explained recall, reflection and potential internal conflict in self image in terms of the distinction between episodic and semantic memory ([21], pages 61-64). Lastly,

[6] This is a 1969 description which presages Dennett's 1995 description of Gregorian Minds. The similarity may not be coincidental. Dennett termed 'Gregorian Minds' after Richard Gregory. The first author of this paper met Richard Gregory and in a subsequent email exchange, in February 2010, Richard Gregory wrote: *" I spent exactly one whole day with John Bowlby when he came to Bristol to visit me and we had a really great day. For most of the time we talked about what we were doing in the Brain and Perception laboratory in the Medical School but we did also talk about his work and I have spent quite a lot of time reading his papers and commenting on them. He did indeed have very wide interests and did think in terms of cybernetics and especially interacting machines. I must say I liked him very much indeed, and he certainly had a sharp and imaginative mind."*

but of particular pertinence to computational modelling, Bowlby describes how much of our habitual processing is automatic and portrays the kinds of reflective meta-processing on mental life which occurs in therapy (at least when therapy is successful) in explicitly computational terms:

> *"The psychological state may then be likened to that of a computer that, once programmed, produces its results automatically when activated. Provided the programme is the one required, all is well.* [When] *representational models and programmes are well adapted, the fact that they are drawn on automatically and without awareness is a great advantage. When however, they are not well adapted, for whatever reason, the disadvantages of the arrangement become serious.*
>
> *For the task is of changing an over-learned programme of action and/or of appraisal is enormously exacerbated when rules long implemented by the evaluative system forbid its being reviewed.* [...] *A psychological state of this kind in which a ban on reviewing models and action systems is effected outside awareness is one encountered frequently during psychotherapy. It indicates the existence of another stage of processing at which defensive exclusion can also take place, different to the stage at which perceptual defence takes place.* ([21], p. 55-56)

What Bowlby is referring to when he discusses therapy are meta-processes such as self-reflection and meta-management [68]. Self-reflection is a less precise term and is involves an organisms or software agent reasoning about itself [54]. The lack of precision is in part because contemporary cognitive science does not have a clear theory of phenomena posited at the 'personal' (as opposed to 'sub-personal') level of description [27]. Beaudoin defines meta-management more precisely as *managing management processes (some of which might be meta-management processes"*. For example meta-management processes control management processes by deciding when to make decisions (such as adopting a goal), and to decide which management processes to activate (such as managing by ruminating or managing by reacting with a well practiced and automatised routine (such as count to ten before responding). [68] provides a recent review of meta-management in the context of meta-cognition. However, this remains an under-explored area of artificial intelligence and cognitive systems research ([74], p. 12).

3 CONCLUSION

Attachment Theory originated from theories set out by Freud, Klein and other psychoanalysts. However, Bowlby ultimately formed a clear distinction between Attachment Theory and Psychoanalytic Theory by explaining the richness, broad scope and complexity of behavioural phenomena which interest psychoanalysts in information processing terms which are congruent with contemporary cognitive science constructs. Bowlby's last publication was a historical review of Attachment Theory, co-authored with Mary Ainsworth [2], and published posthumously in 1991 (a year after Bowlby's death in 1990). This historical review concentrates more on the early and medium term history of Attachment Theory rather than developments occurring around the time of its publication. However, in 1990 Emde [43] published a short review of developments in attachment research in that period, which he characterised as *'The Third Phase Of Attachment Research'*. In his view, Bowlby's departure from the British Object Relations School of psychoanalysis was a first phase of attachment research. Mary Ainsworth and co-workers move to study individual differences in infant attachment constitutes a second phase which revitalised attachment research because it helped link theory to individual experience and clinical concepts. Emde proposed a third phase of attachment research was occurring at the time he wrote. It involved attachment research moving beyond infancy to observe and measure individual differences in the pre-school years. To do this new assessment approaches were created to move away from concentration on the Strange Situation Procedure [1] and address more complex behaviours of toddlers and pre-schoolers.

Since the 1990s the variety and scope of approaches has greatly broadened [25]. Contemporary directions in Attachment Theory include a deepening of areas which were already of interest to Bowlby, including: the caregiving system [33]; evolutionary [65]; and psychopathological [28] perspectives. New directions for Attachment Theory with regard to the behavioural phenomena of interest to Attachment Theory researchers include: romantic relationship [34]; comparative [72]; and cross-cultural [46] perspectives. The scope of Attachment Theory has broadened so that some contemporary research can be viewed as of a type with contemporary social psychology [40] and other research fits within the area of cognitive psychology [31].

Of particular interest to computational emotion modellers is research relevant to understanding how the attachment control system is implemented neurally and physiologically. This includes biological [39] and imaging research on the neuroscience of attachment [24]. A question for computational modellers is whether these contemporary perspectives replace or complement the particular attachment control system approach that Bowlby set out. In addition, whilst this review has demonstrated that Bowlby's attachment control system construct was a clear departure from psychoanalytic structures and mechanisms, researchers of a psychoanalytic persuasion still make a contribution to attachment research. For example, Fonagy suggests that from a psychoanalytic perspective, Attachment Theory and Psychoanalysis are heading towards the same end point, with regard to the phenomena they are trying to explain [35]. Fonagy suggests that from this psychoanalytic perspective Attachment Theory can seem 'method-bound' because of the strong focus on a set of validated measures:

> [Attachment Theory's] *"scope was determined less by what fell within the domain defined by relationship phenomena involving a caretaking-dependent dyad and more by the range of groups and behaviors to which the preferred mode of observation, the strange situation, the adult attachment interview, and so forth, could be productively applied. This sheltered the theory from a range of ideas that clinical psychoanalysts evolved, particularly in the context of analytic work with increasingly severely disturbed chronic personality-disordered individuals."*([35], pp 472-473).

Therefore, Fonagy suggests that Attachment Theory might benefit from engaging with clinical discoveries of psychoanalysts. Psychoanalysts have also taken issue with the explanatory power of Bowlby's attachment control system constructs [35]. For example, the psychoanalyst Jeremy Holmes suggests that Attachment Theory lacks complexity because it provides less focus on emotional phantasy about loss ([42] pp 6-7). To sum up Attachment Theory research since Bowlby, whilst contemporary Attachment Theory is diverse and multifaceted with many varied perspectives, most of these perspectives involve collecting new kinds of empirical data not new underlying information processing explanations for Attachment Theory. The jury is still out whether new biological and imaging results

will replace or complement Bowlby's psychological level formulation of the attachment control system.

Petters and Waters [58] provide a review of recent attempts in computational attachment modelling. These include autonomous software agent models, robotic models, multi-generational evolutionary models and artificial neural network models. However, when this range of studies is compared to the breadth and depth of phenomena that Bowlby discussed it is apparent that taken together the collection of all existing attachment models is rather incomplete in two senses. Firstly, this collection provides just a few examples of possible models of the attachment control system. So the relatively small number of extant attachment models gives a sparse coverage of possible models. Secondly, the simulations that have been created are relatively simple compared to Bowlby's idea of a complex hierarchy of control systems ([21], p. 52).

A key direction for future work in a computational modelling approach to attachment is to systematically explore the architectural design space of possible attachment control systems. What are the options a modeller might choose from in designing cognitive architectures? One way to classify different architectures is by the nature of the subsystems they possess. For very simple architectures, with few subsystems in uncomplicated arrangements, this might be a reasonable approach. For example how infants respond in separation and reunion episodes might be simulated using predictive processing [55]; or the verbal interchange between adults discussing attachment issues might be simulated using general cognitive principles such as frequency and decay memory functions rather than specific language processing mechanisms [56]. However, for complex architectures it may be that defining architectures by their component parts may make no sense. For architectures which are sufficiently rich, with numerous independent components, functional differentiation of components and a variety of causal linkages, how the components are arranged will be highly influential in determining the architecture's capabilities [67, 76]. It may even be that when architectural components have some minimal sufficient structural variability and speed of operation, then the arrangement of the components becomes crucial. This idea phrased more succinctly is that *"architecture dominates mechanism - global design normally determines global capabilities to a greater extent than implementation details"* ([76], p.104). In the context of attachment modelling: how the attachment control system's components are arranged may matter more than the performance details of each component.

Dennett portrays two contrasting approaches to pursuing a deeper understanding of the information processing architectures: *"At one pole, are AI visionaries who don't write actual programs but settle instead for facts about specs for any successful program for one task or another, and at the other pole, hard bitten engineering types who aren't impressed until they see code that actually runs and does the job. The rest, they sneer, is not software, it is vapourware"* ([30], p. 139). Bowlby's approach was clearly of the 'visionary' variety. Much contemporary emotion modelling is 'hard bitten' engineering in the sense that it is strongly constrained by the requirement to produce working simulations, even if these are relatively narrow in scope or application. The challenge for the field of attachment modelling is whilst producing running simulations to not lose site of Bowlby's complex vision of the overall attachment control system and the richness of phenomena it explains. A key conclusion for this historical review is the value of the rich, deep, broad and complex modelling scenarios that the attachment domain provides. Empirical attachment data cover the lifespan of social and emotional development, in caregiving and romantic relationships. Attachment is also researched from evolutionary, comparative, biological, cross-cultural, clinical, emotion-regulation, and educational perspectives [25], providing a storehouse of constraints for computational attachment models.

4 CODA - A COUNTERFACTUAL ALTERNATIVE HISTORY

After concluding a historical account of the development of the attachment control system in the section above, this final section of the paper will do something different by speculating about how Attachment Theory might have developed if history had gone differently. The systemic family therapy practitioner Robin Routledge [62] describes Bowlby's (1949) paper: *'The Study and Reduction of Group Tensions in the Family'* [8] as possessing a clear systemic perspective and approach. It therefore predates other examples of systemic analysis and systemic family therapy- being written in the 1940s at a time of scientific origin for many branches of modern psychotherapy. Routledge notes that whilst the Attachment Theory that was developed subsequently to this paper was an extension away from Object Relations Psychodynamic Theory, it was actually an even bigger step away from the more systemic view found in the 1949 paper. Routledge quotes John Byng-Hall's recollection that Bowlby took this decision because he believed the infant-mother dyad *"was complex enough itself"* (Bing-Hall quoted in [62], p 17) and needed to be explained before the broader family system was investigated. This leaves Routledge pondering 'what might have been' in attachment research if Bowlby had made a different decision at this time. Therefore a final thought in this alternative history is to imagine a counter factual alternative history for Attachment Theory, more strongly focused on dyads embedded in broader systems. Computational techniques for modelling emotion with multi-agent systems might spearhead this sort of integrated research today - using existing attachment models to simulate wider family systems.

REFERENCES

[1] M. Ainsworth, M. Blehar, E. Waters, and S. Wall, *Patterns of Attachment: A psychological study of the strange situation*, Erlbaum, Hillsdale, NJ, 1978.

[2] M. Ainsworth and J. Bowlby, 'An ethological approach to personality development', *American Psychologist*, **46**, 333–341, (1991).

[3] J.R. Anderson, *How Can the Human Mind Occur in the Physical Universe?*, OUP, New York, 2009.

[4] C.G. Bell and A. Newell, *Computer structures: readings and examples*, McGraw-Hill Book Company, New York, 1971.

[5] J. Bowlby, 'The influence of early environment in the development of neurosis and neurotic character.', *International Journal of Psychoanalysis*, **21**, 154–178, (1940).

[6] J. Bowlby, 'A parent at hospital. letters to the editor.', *The Lancet*, **June 2**, 704, (1940).

[7] J. Bowlby, 'Forty-four juvenile thieves: Their character and home life.', *International Journal of Psychoanalysis*, **25**, 1–57, (1944).

[8] J. Bowlby, 'The study and reduction of group tensions in the family', *Human Relations*, **2**, 123–8, (1949).

[9] J. Bowlby, *Maternal Care and Mental Health*, World Health Organisation WHO, New York, 1951.

[10] J. Bowlby, 'Psychoanalytic instinct theory.', in Discussions on Child Development: Volume 1, *eds. J. M. Tanner and B. Inhelder*, 182–187, Tavistock Publications, London, (1956).

[11] J. Bowlby, 'The nature of a child's tie to his mother', *International Journal of Psychoanalysis*, **39**, 350–373, (1958).

[12] J. Bowlby, 'Grief and mourning in infancy and early childhood', *The psychoanalytic study of the child*, **XV**, 9–52, (1960).

[13] J. Bowlby, 'Processes of mourning', *International Journal of Psycho-Analysis*, **42**, 317–340, (1960).

[14] J. Bowlby, 'Separation anxiety', *International Journal of Psycho-Analysis*, **41**, 89–113, (1960).

[15] J. Bowlby, 'Symposium on psycho-analysis and ethology ii: Ethology and the development of object relations.', *International Journal of Psycho-Analysis*, **41**, 313–317, (1960).

[16] J. Bowlby, 'Separation anxiety: A critical review of the literature', *Journal of Child Psychology and Psychiatry*, **1**, 251–269, (1961a).

[17] J. Bowlby, 'Pathological mourning and childhood mourning', *Journal of the American Psychoanalytic Association*, **11**, 500–541, (1963).

[18] J. Bowlby, *Attachment and loss, Volume 1: Attachment*, Basic books, New York, 1969.

[19] J. Bowlby, *Attachment and loss, Volume 1: Attachment*, Basic books, New York, 1969 | 1982. (Second edition 1982).

[20] J. Bowlby, *Attachment and loss, Volume 2: Separation, anxiety and anger*, Basic books, New York, 1973.

[21] J. Bowlby, *Attachment and loss, volume 3: Loss, sadness and depression*, Basic books, New York, 1980.

[22] J. Bowlby, 'Contribution to symposium, 'Emanuel Peterfreund on information and systems theory", *he Psychoanalytic Review*, **68**, 187–90, (1981).

[23] I. Bretherton, 'The origins of attachment theory: John Bowlby and Mary Ainsworth', *Developmental Psychology*, **28**(5), 759–775, (1992).

[24] I. Bretherton and K.A. Munholland, 'The internal working model construct in light of contemporary neuroimaging research', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 63–91, Guilford Press, London, (2016).

[25] J. Cassidy and P.R. Shaver, Handbook of Attachment, 3rd Ed., Guilford Press, London, 2016.

[26] J.M. Chappell and A. Sloman, 'Natural and artificial meta-configural altricial information-processing systems', *International Journal of Unconventional Computing*, **3(3)**, 211–239, (2007).

[27] T. Deacon, *Incomplete Nature: How Mind Emerges from Matter*, W. W. Norton and Company, London, 2012.

[28] M. Deklyen and M.T. Greenberg, 'Attachment and psychopathology in childhood', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 639–667, Guilford Press, London, (2016).

[29] D. C. Dennett, *Brainstorms: Philosophical Essays on Mind and Psychology*, MIT Press, Cambridge, MA, 1981.

[30] D.C. Dennett, *Intuition Pumps and Other Tools for Thinking*, Penguin Press, London and New York, 2013.

[31] M.J. Dykas and J. Cassidy, 'Attachment and the processing of social information across the life span: Theory and Evidence', *Psychological Bulletin*, **137**(1), 19–46, (2011).

[32] A Edalat and F Mancinelli, 'Strong attractors of Hopfield neural networks to model attachment types and behavioural patterns'. IEEE, (2013).

[33] B.C. Feeney and S.S.Woodhouse, 'Caregiving', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 827–851, Guilford Press, London, (2016).

[34] J.A. Feeney, 'Adult romantic attachment: Developments in the study of couple relationships', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 827–851, Guilford Press, London, (2016).

[35] P. Fonagy, 'Points of contact and divergence between psychoanalytic and attachment theories: Is psychoanalytic theory truly different.', *Psychoanalytic Inquiry*, **19:4**, 448–480, (1999).

[36] R.C. Fraley, 'A connectionist approach to the organization and continuity of working models of attachment', *Personality and Social Psychology Review*, **6**, 1157–80, (2007).

[37] Nico H. Frijda, *The emotions*, Cambridge University Press, Cambridge, 1986.

[38] W. Grey Walter, 'Presentation: Dr Grey Walter.', in Discussions on Child Development: Volume 2*, eds. J. M. Tanner and B. Inhelder*, 21–74, Tavistock Publications, London, (1956).

[39] A.A. Hane and N.A. Fox, 'Studying the biology of human attachment', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 223–242, Guilford Press, London, (2016).

[40] C. Hazan and P. Shaver, 'Attachment as an organizational framework for research on close relationships', *Psychological Inquiry*, **5**, (1994).

[41] R.A. Hinde, 'Ethology and child development', in *Hand book of child psychology, eds. P.H. Mussen*, 27–93, J. Wiley and Sons, New York, (1983).

[42] J. Holmes, *John Bowlby and Attachment Theory*, Routledge, 1993. (revised edition).

[43] M. Main and J. Solomon, 'Preface: The third phase of attachment research', in Attachment in the Preschool Years*, eds. M.T. Greenberg and D. Cicchetti and E.M. Cummings*, ix–xii, The University of Chicago Press, Chicago, (1990).

[44] S. Marsella and J. Gratch, 'Computationally modeling human emotion', *Communications of the ACM*, **57**(12), 56–69, (2014).

[45] S. Marsella, J. Gratch, and P. Petta, 'Computational models of emotion', in *A Blueprint for Affective Computing-A sourcebook and manual*, eds., K. Scherer, T. Banziger, and E. Roesch, 21–46, Oxford University Press, Oxford, (2010).

[46] J. Mesman, M.H. van Ijzendoorn, and A. Sagi, 'Cross-cultural patterns of attachment: Universal and contextual dimensions', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 852–878, Guilford Press, London, (2016).

[47] U. Neisser, 'The imitation of man by machine', *Science*, **139**, 193–197, (1963).

[48] A. Newell, *Unified Theories of Cognition*, Harvard University Press, Cambridge, MA, 1990.

[49] D. Petters, 'Simulating infant-carer relationship dynamics', in *Proc AAAI Spring Symposium 2004: Architectures for Modeling Emotion - Cross-Disciplinary Foundations*, number SS-04-02 in AAAI Technical reports, pp. 114–122, Menlo Park, CA, (2004).

[50] D. Petters, 'Building agents to understand infant attachment behaviour', in *Proceedings of Modelling Natural Action Selection*, eds., J.J. Bryson, T.J. Prescott, and A.K. Seth, 158–165, AISB Press, School of Science and Technology, University of Sussex, Brighton, (2005).

[51] D. Petters, *Designing Agents to Understand Infants*, Ph.D. dissertation, School of Computer Science, The University of Birmingham, 2006. (Available online at http://www.cs.bham.ac.uk/research/cogaff/).

[52] D. Petters, 'Implementing a theory of attachment: A simulation of the strange situation with autonomous agents', in *Proceedings of the Seventh International Conference on Cognitive Modelling*, 226–231, Edizioni Golardiche, Trieste, (2006).

[53] D. Petters, 'Bringing Together Different Pieces to Better Understand Whole Minds', in *From Robots to Humans and Back: Reflections on Hard Problems in the Study of Cognition.*, eds., J.L. Wyatt, D. Petters, and D. Hogg, (2014).

[54] D. Petters, 'Losing Control in the H-CogAff Architecture', in *From Robots to Humans and Back: Reflections on Hard Problems in the Study of Cognition.*, eds., J.L. Wyatt, D. Petters, and D. Hogg, (2014).

[55] D. Petters, 'An encounter between 4e cognition and attachment theory', *Connection Science*, **28**(4), 387–409, (2016).

[56] D. Petters and L. Beaudoin, 'Attachment Modelling: From Observations to scenarios to designs', in *Computational Neurology and Psychiatry: Volume 6 of Springer Series in Bio-/Neuroinformatics.*, eds., P. Erdi, B.S. Bhattacharya, and A. Cochran, 227–271, (2017).

[57] D. Petters and E. Waters, 'A.I., Attachment Theory, and simulating secure base behaviour: Dr. Bowlby meet the Reverend Bayes', in *Proceedings of the International Symposium on 'AI-Inspired Biology', AISB Convention 2010*, 51–58, AISB Press, University of Sussex, Brighton, (2010).

[58] D. Petters and E. Waters, 'Modelling Emotional Attachment: An Integrative Framework for Architectures and Scenarios', in *Proceedings of IJCNN*. IEEE, (2015).

[59] D. Petters, E. Waters, and F. Schönbrodt, 'Strange carers: Robots as attachment figures and aids to parenting', *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, **11**(2), 246–252, (2010).

[60] R.W. Picard, *Affective Computing*, MIT Press, Cambridge, Mass, London, England, 1997.

[61] M. Richardson, K. Shockley, B. Fajen, M. Riley, and M. Turvey, 'Ecological Psychology: Six principles for an embodied embedded approach to behavior.', in *Handbook of Cognitive Science: An Embodied Approach*, eds., P. Calvo and T. Gomila, 161–187, Elsevier, San Diego, CA, (2010).

[62] R. Routledge, 'The first reported family therapy meeting?', *Context*, **115**, 18–24, (2011).

[63] K.R. Scherer, 'Emotion and emotional competence: conceptual and theoretical issues for modelling agents.', (2010).

[64] H. A. Simon. Motivational and emotional controls of cognition, 1967.

[65] J.A. Simpson and J. Belsky, 'Attachment theory within a modern evolutionary framework', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 91–117, Guilford Press, London, (2016).

[66] A. Sloman, 'Towards a grammar of emotions', *New Universities Quarterly*, **36**(3), 230–238, (1982).

[67] A. Sloman, 'The mind as a control system', in *Philosophy and the Cognitive Sciences*, eds., C. Hookway and D. Peterson, 69–110, Cambridge University Press, Cambridge, UK, (1993).
[68] A. Sloman, 'Varieties of meta-cognition in natural and artificial systems', in *Metareasoning: Thinking about thinking*, eds., M. T. Cox and A. Raja, 307–323, MIT Press, Cambridge, MA, (2011). (Available at http://www.cs.bham.ac.uk/research/cogaff/).
[69] A. Sloman and M. Croucher, 'Why robots will have emotions', in *Proc 7th Int. Joint Conference on AI*, pp. 197–202, Vancouver, (1981).
[70] T.S. Smith, G. Stevens, and S. Caldwell, 'The familiar and the strange: Hopfield network models for prototype-entrained attachment-mediated neurophysiology', in *Mind, brain, and society: Toward a neurosociology of emotion, Vol. 5.*, eds., D.D. Franks and T.S. Smith, 213–245, Elsevier Science/JAI Press, (1999).
[71] G. Sommerhof, *Analytical Biology*, Oxford University Press, London, 1950.
[72] S.J. Suomi, 'Attachment in rhesus monkeys', in Handbook of Attachment, (Third edition *, eds. J. Cassidy & P.R. Shaver*, 133–155, Guilford Press, London, (2016).
[73] F. van der Horst, *John Bowlby - From Psychoanalysis to Ethology: Unravelling the Roots of Attachment Theory*, Wiley-Blackwell, Chichester, 2011.
[74] D. Vernon, *Artificial Cognitive Systems: A Primer*, MIT Press, Cambridge: MA, 2014.
[75] E. Waters, K. Kondo-Ikemura, G. Posada, and J. Richters, 'Learning to love: Mechanisms and milestones', in Minnesota Symposium on Child Psychology (Vol. 23: Self Processes and Development)*, eds. M. Gunner & Alan Sroufe*, 217–255, Psychology Press, Florence, KY, (1991).
[76] I.P. Wright, A. Sloman, and L.P. Beaudoin, 'Towards a design-based analysis of emotional episodes', *Philosophy Psychiatry and Psychology*, **3**(2), 101–126, (1996).