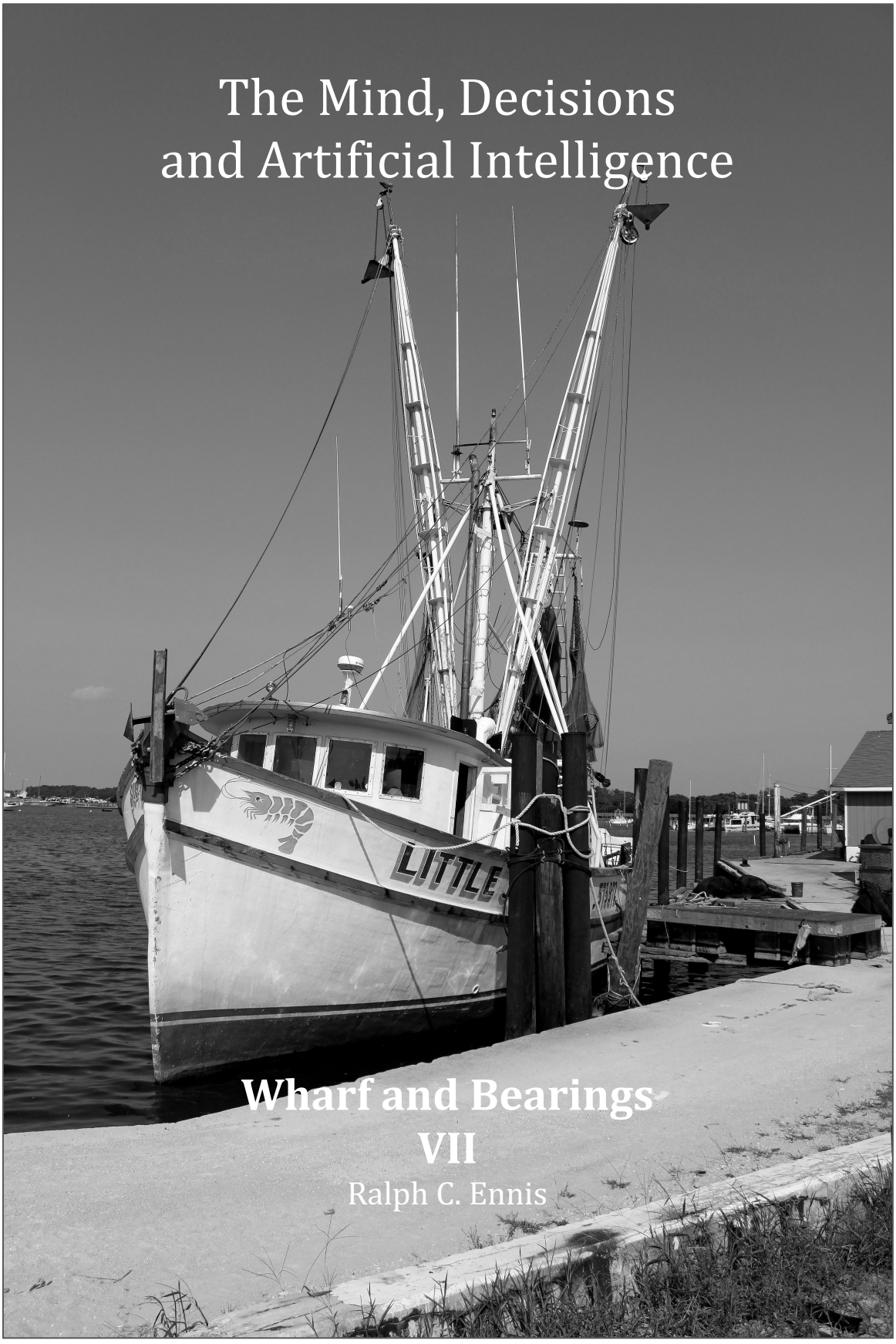
The Mind, Decisions and Artificial Intelligence



Wharf and Bearings
VII

Ralph C. Ennis

Wharf and Bearings VII

Essays on
The Mind, Decisions and Artificial Intelligence

Ralph C. Ennis

Wharf and Bearings VII

The Mind, Decisions and Artificial Intelligence

Table of Contents

Introduction to WB Series 4
A Model of the Mind 6
Assumptions and Pathway for Achieving Artificial General Intelligence 26
Ethical DNA Model for Artificial General Intelligence 35
Evaluating Virtue and Vice Values for Ethical Artificial General Intelligence 46
Thought Dynamo Decision Mapping Model 54
A Theoretical Model for Research in Intercultural Decision Making 85
Resources to Explore 88
About the Authors 89

Introduction to WB Series

As land dwelling beings, when we are at sea, two thoughts are always in mind—even if in the back of our minds. Where's the wharf—a safe place to dock? What's my bearing—which direction am I going and will it in time get me to a safe wharf?

We all need safe places. Fundamentally, life is filled with uncertainties. Sometimes we feel in control or at minimal risk of danger, but that is always only temporary. Our common human experience inevitably involves risks of the unknown. And through death, not to be feared, we journey to an eternal wharf in Jesus.

We all need a bearing that will get us where we want to go while preserving the ability to plot a course to a safe place. We can play far from this wharf, but we desire home.

Wharf and Bearings Series, a collection of essays, poems and a short story, is intended to present a journey—my journey. I share it with the hope that it will give some guidance as you seek your bearings and wharf throughout your life.

At no time are these essays to be considered exhaustive, they are pathways I have taken to find wharf and bearings for me. And collectively, they represent "philosophical peace" for me.

I dedicate this series to my wife of 45 years (in 2018), our four children and their spouses and our 15 grandchildren and the generations to follow! Here's an overview of the series:

Wharf and Bearings ONE:
Hope and Beholding the Triune God

Wharf and Bearings TWO:
Spirituality and the Triune God

Wharf and Bearings THREE:
Knowing, Beauty, Ethics and Reality

Wharf and Bearings FOUR:
Love, Longings, Success and Consequences

Wharf and Bearings FIVE:
Poems from the Soul

Wharf and Bearing SIX:
Oneness in Marriage

Wharf and Bearings SEVEN:
The Mind, Decisions and Artificial Intelligence

Wharf and Bearing EIGHT:
World View and Culture

Wharf and Bearing NINE:
Gospel Implications

Wharf and Bearing TEN:
Our Times and Futures

"… we will tell the next generation the praiseworthy deeds of the LORD, his power, and the wonders He has done …"
Psalm 78:4 -6

A Model of the Mind

From an Embedded Infinities Perspective

Table of Contents

Abstract 7

Part One: Embedded Infinities as a Model of the Mind 8
- Embedded Infinities – A Master Construct 8
- Mental Space and Outside Space 8
- Decision Making – the Processing of the Mind 9
- Holographic Neural Networks and Behavioral Extensions 16
- Probabilistic Cause and Effect Embedded in Time 16
- Metaphysical Thought – the 'Beyond' of Embedded Infinities 17
- Reliability of Knowing the Mind and Beyond 17

Part Two: Brief Applications for a Model of the Mind 17
- Global Economic Theory 18
- Market Research Strategies 18
- Philosophy 18
- Metaphysics 19
- Psychological Health Theory 19
- Educational Agenda and Ethics 20
- Societal Health Theory 21
- Intercultural Communication and Relations 21
- International Conflict Negotiations 21
- Physics 22
- Biology 22
- Artificial Intelligence Programming 23

Bibliography 24

April 26, 2008

Abstract

In an effort to model the human mind, one must chose a place to begin. The utmost abstraction conceived by the human mind is posited as a starting point. From that perspective the end (of thought) can be perceived from the beginning and the beginning from the end.

The concept of embedded infinities is put forth as the highest abstraction of human thought. Embedded infinities allows for total oneness and complete discrete representation (with or without order) and all in-between. Thus, all solutions possible for the mind to conceive with congruence, contradiction, paradox, comedy or disillusionment are implicit within the framework of the abstraction of embedded infinities. It is a central construct by which the mind can be explored by the mind.

In order to make a case for embedded infinities as a master construct of the mind several arenas must be considered: 1) the **nature of space** (as examined and explained by the mind), 2) the **nature of decision making** (the core processing of the mind) from a multicultural perspective, 3) **holographic neural networks**, 4) **probabilistic cause and effect embedded in time**, 5) the **nature of metaphysical thought** (the beyond of embedded infinities) and 6) the **reliability of knowing** the mind and beyond. In other words, a culture-general model of the mind must account for all possible knowledge of the mind in the physical, psycho-social, biological and metaphysical worlds and account for the reliability of knowing across time.

From these six perspectives, a model of the mind can be posited – all embracing embedded infinities as a master abstract construct. From the decision making model a mathematical model and fuzzy optimization of the mind will be suggested. This model will rely on the mind's conceptualization of space. Finally, a wide range of applications and research agendas will be suggested.

Part One: Embedded Infinities as a Model of the Mind

Embedded Infinities – A Master Construct

The concept of embedded infinities suggests that which is discrete but which intersects and is intersected by in an infinite manner.

The summation of embedded infinities is universal oneness with a beyond. Embedded discreteness allows for complete uniformity of discrete elements. And all variations of embedded discrete elements allows for all solutions of diversity – both physical and mental and into the metaphysical. Embedded infinities allows for total oneness and total diversity and all in-between and thus all solutions of the mind.

Are embedded infinities the purview of only a few or is this abstraction a culture-general phenomenon of which both children and adults are aware? The terminology of embedded infinities is familiar to few. The concept, however, is posited as universal. A child's quest for understanding what's beyond a place he or she cannot see, the desire to explain death in terms of life, the romance of a another that merges the emotional, sexual worlds of two people, dichotomous and continuous moral categories, resurrection and reincarnation, the mathematical inquiry of a greatest prime number, the set of even numbers embedded within the set of real numbers – all these are rooted in the abstraction of embedded infinities. These are common pursuits across many cultures.

A simplistic illustration of embedded infinities is a granite block at the hands of a master sculpturer. The artist sees an infinite number of possibilities embedded within a single block – and regardless the shape of his work, infinite possibilities remain.

This model of the mind – that all thinking is embedded within the abstraction of embedded infinities – is based on the assumption that the mind is incapable of thinking outside itself and yet the mind perceives a beyond apart from and yet connected to itself.

Mental Space and Outside Space

The mind is incapable of thinking outside itself and relies on embeddedness to relate to a perceived outside – a beyond the boundary of the mind. An individual mind is embedded in a physical world and a social world and into a metaphysical one that it perceives as real or imagined or even non-existent.

The pursuit to explain a universe of space can be reconstructed into the pursuit to explain mental space since space can only be mentally perceived (as it interprets sensory input). The final quest for real, imagined and/or metaphysical space can only lie within the boundaries of mental space. *Mental space has as it's boundary embedded infinities which allows for total oneness and complete discrete representation and all variations in-between and thus all solutions possible for the mind to conceive including all mental solution available to the nature of space.*

Mental modeling of space (including n^{th} dimensional modeling and multiple universes), therefore, lies somewhere within the concept of embedded infinities since any and all mental models are limited to the abstraction boundaries of the mind. Since all mental solutions (and the mind is limited to itself)

are conceived within embedded infinities, which mental solution best describes the nature of a universe of space? Or posed differently, if space is the discrete that formulates all else (time, energy, matter) and thus contains "all solutions", then is space not best represented by embedded infinities? The answer is not necessarily yes, but the question is valid to pursue. And so this section is in pursuit of modeling external space of all physical solutions from the perspective of embedded infinites – the all solution of mental space.

From space all sensory input is derived. For example, the modulation of sound waves is fundamentally a variation of spatial properties through time, and in like manner so are all sight, touch, smell and taste. Embedded space is the extreme abstraction for categorizing a universe to which humans are limited, by their sensory faculties and imaginations, to perceive.

A model of the mind must account for all possible knowledge and imaginations of the mind (across time) in the physical, psycho-social, biological and metaphysical worlds. If space is conceived to be the fundamental reality of the physical world, then a model of the mind must account for all mental modeling of physical space.

Modeling Physical Space from the Perspective of Mental Space

Physical space (that which is outside the mind) has long been geometrically modeled from a 3-dimensional perspective and as continuous or discrete space with an additional time dimension. Quantum models have suggested more dimensions and even infinite dimensions. The origin of space has been conceptualized from observable data as beginning from a Big Bang or multiple Big Bangs and thus the possibility of an infinite number of universes. All of these models fall within the limitations of the mind as an unit that generates and plays with embedded infinities abstractions.

Though there are an infinite number of possibilities for modeling physical space, to find a reliable one is to find one that most accurately describes the sensory observable data available. That subject will be addressed in a future paper.

Decision Making – the Processing of the Mind

Decision Making as the Foundational Process of the Mind

The central process of the mind can be posited as decision making. This process involves sensory spatial input, categorization, fuzzy-step boundaries of consciousness, discrete (bounded continua of) memory (storage, weighing, modification and retrieval), emotion, intellect and imagination. All that the mind does can be conceptualized as making decisions continuously at various levels of consciousness and with various levels of confidence. The mosaic of these discrete decisions can be modeled as embedded infinities of decisions.

Each decision is reinforced and/or challenged by new decisions. Thoughts are decisions of the mind weighted with feelings of sensory inputs that have been interpreted into emotions. The miracle of thoughts (verbal decisions to represent spatial variations – i.e. varying levels of abstractions of

spatial variations), of emotions wrapped onto thought, and of fuzzy categories of thought are processed through mental decision making (at various levels of consciousness) that has memory of a past presences and imagines a future presence and uses external symbols as memory props (i.e. conscious shifters).

Can emotions be in concert with mental abstractions? Feelings (of pain and pleasure, often linked to variations of energy levels) are viewed as panhuman responses to sensory inputs of spatial variations. Emotions are culturally constructed abstractions of those feelings. To emotionalize embedded infinities is to come face to face with jealousy – varying degrees of possessiveness of (full, partially occupied or null) space. Thus, jealousy is posited as a central tendency of all emotions and is rooted in the abstraction limitation of the mind that is occupied with managing in the present, jealous, embedded infinities of space while seeking a creative harmonious future with a view to the past.

Can a culture-general process of decision making be conceptualized from an embedded infinities perspective? If so, will the model be "contrived" or "real" or simply reflect the limitations of the mind to process from the highest level of abstraction it is capable of? What is suggested is the latter. An intercultural decision making (ICDM) model will be constructed in this section. A detailed description of the ICDM may be found in "A Theoretical Model for Research in Intercultural Decision Making" (Ennis, 2002).

A sufficient process model of the mind (i.e. decision making) will account for learning decisions, psychological health and insanity choices, personality, sexual decision making, cultural health, economic systems, organization patterns, beauty, language rhetoric, instinct, antithesis and synthesis, human bonding, mathematics and other diverse topics.

The ICDM model posited will involve three intersecting axes of logic of intellect, logic of emotion and imagined outcomes embedded within an input-feedback system. These intersecting axes can be conceptualized to account for such notions as beauty, passion and love. [The 3-d embedded infinity of interrelated logics and imagined outcomes will be employed in a subsequent paper to construct a mathematical representation of the mind that can be optimized.]

Modeling Decision Making from an Embedded Infinities Perspective

Decision making is wrapped in a complex web of interactive factors. This web can be seen as an infinite loop that is embedded within all thought and decision.

An overall flow of decision making (see figure 1) is proposed in "A Thought Dynamo Model of Decisions" (Ennis, nd). Decision making flows from five types of decisions (plot sensory inputs, association and adjustment decisions, solidification and employment decisions) that are made through a central thought process of decision making. That thought process, referred to as a thought dynamo, involves a logic of intellect, logic of emotion and imagined outcomes (see figures 2).

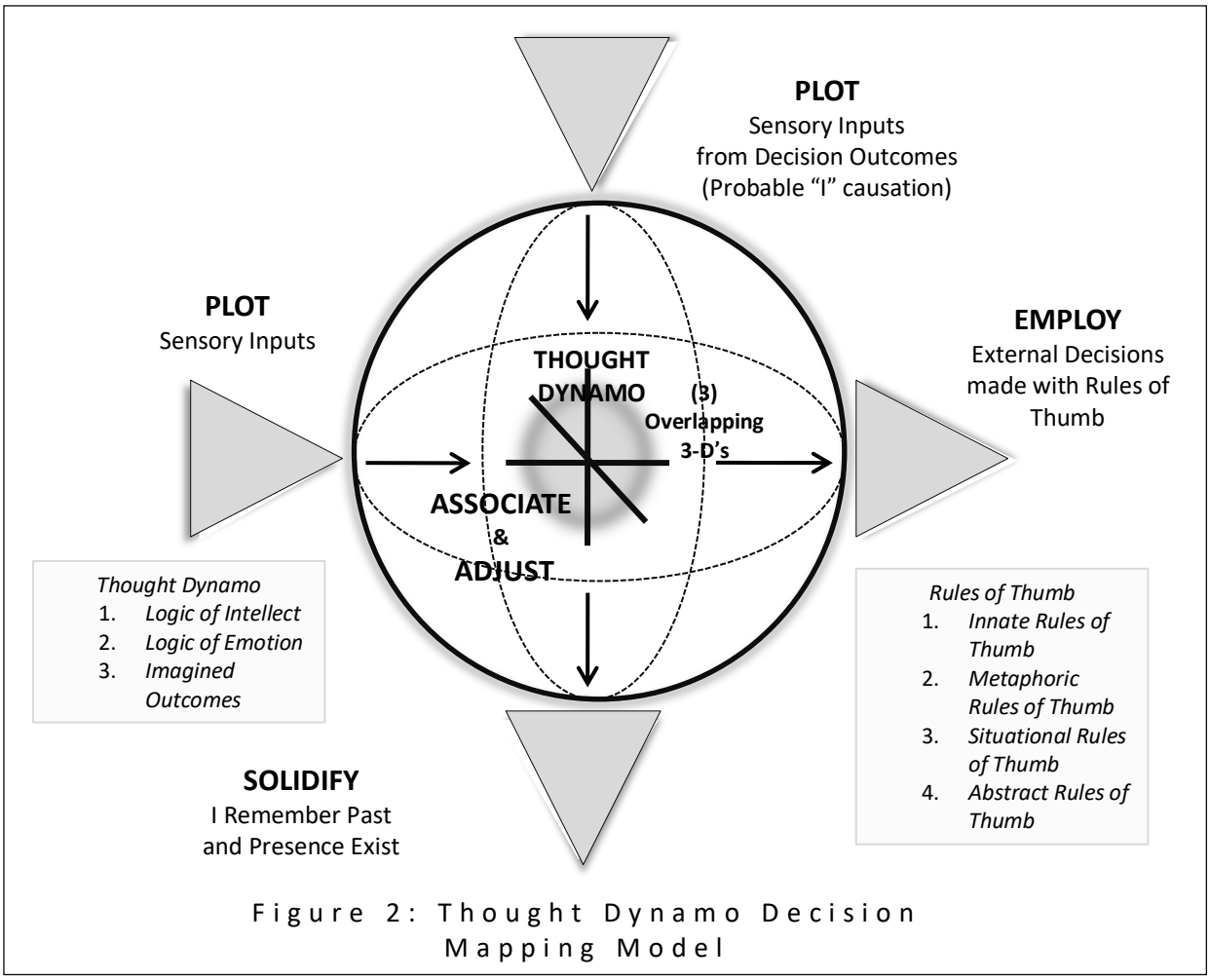


Figure 2: Thought Dynamo Decision Mapping Model

Logic of Intellect

The logic of intellect in this model has three axes: power axis, moral axis, and certainty axis (see figure 2). The power axis is a continuum from powerful to powerless. The moral axis is a continuum from good to evil. The certainty axis is a continuum from accuracy to intuition. And the central construct of the logic of intellect is space. Each of the specifics of the axes is defined by every culture. However, the model posits that all cultural logics involve these three axes of continua.

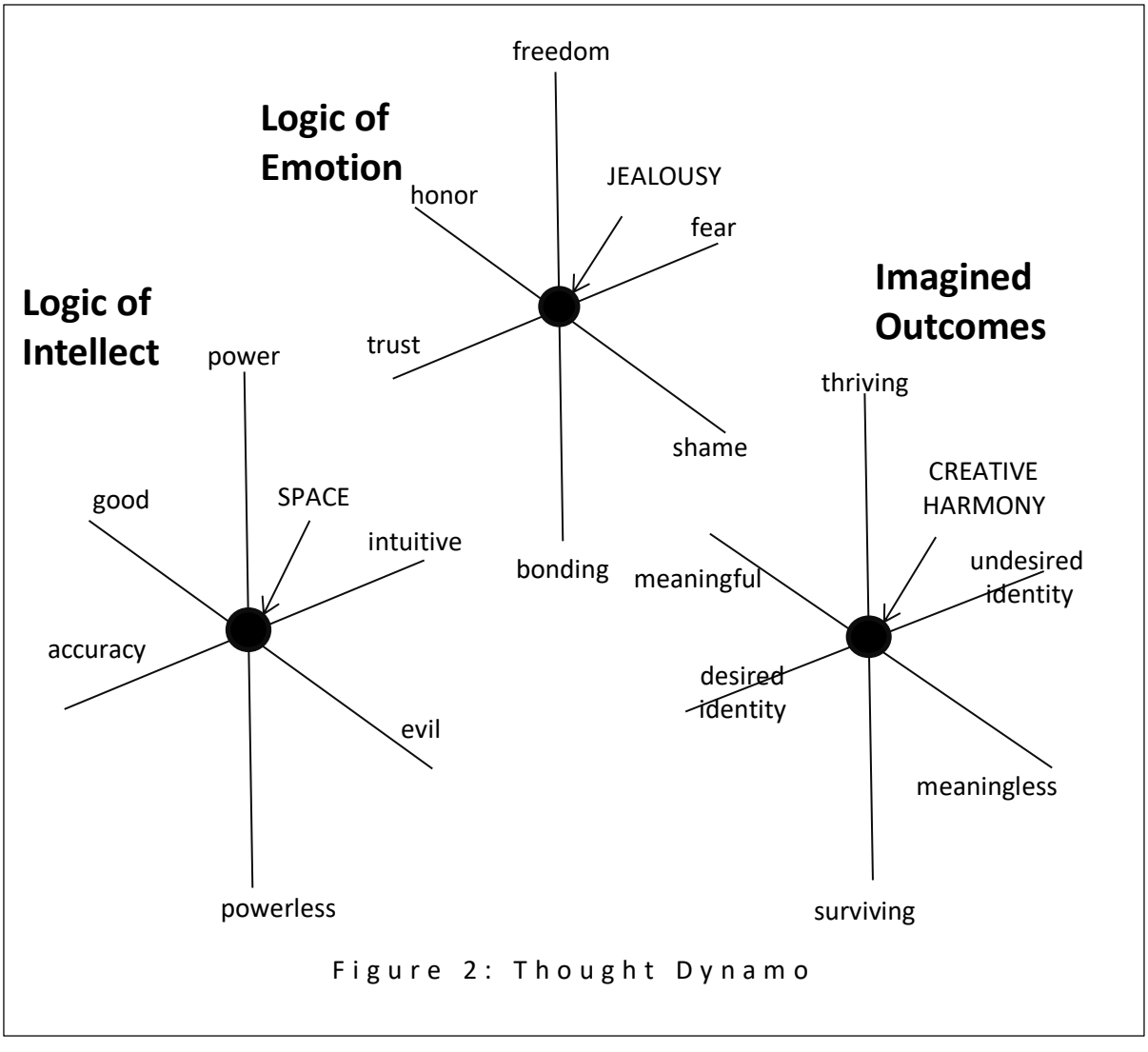


Figure 2: Thought Dynamo

Logic of Emotion

Similarly, the logic of emotion (see figure 2) is posited to involve three axes: relational, hierarchal and liberty. The relational axis is a continuum from trust to fear. The hierarchal axis is a continuum from honor to shame. The liberty axis is a continuum from freedom to bonding. And the central construct of the logic of emotion is jealousy.

Imagined Outcomes

The construct of imagined outcomes in the model is posited to involve three axes: adaptation, meanings and identity. The central construct of imagined outcomes is posited as creative harmony (see figure 2).

As a present process, the logic of intellect and emotions both have two time elements embedded in them. The past has been process and stored through these axes. And a second time element can be conceptualized as imagined outcomes. While acknowledging that the duration of time implicit in imagination is culture-specific, this ICDM model positions imagination into the non-present as a culture-general construct. A future imagined outcomes emphases the possibility of decisions modifying a present assessment of reality. These imagined outcomes create the possibility of re-interpreting the past with present decisions.

Three Intersecting Axes

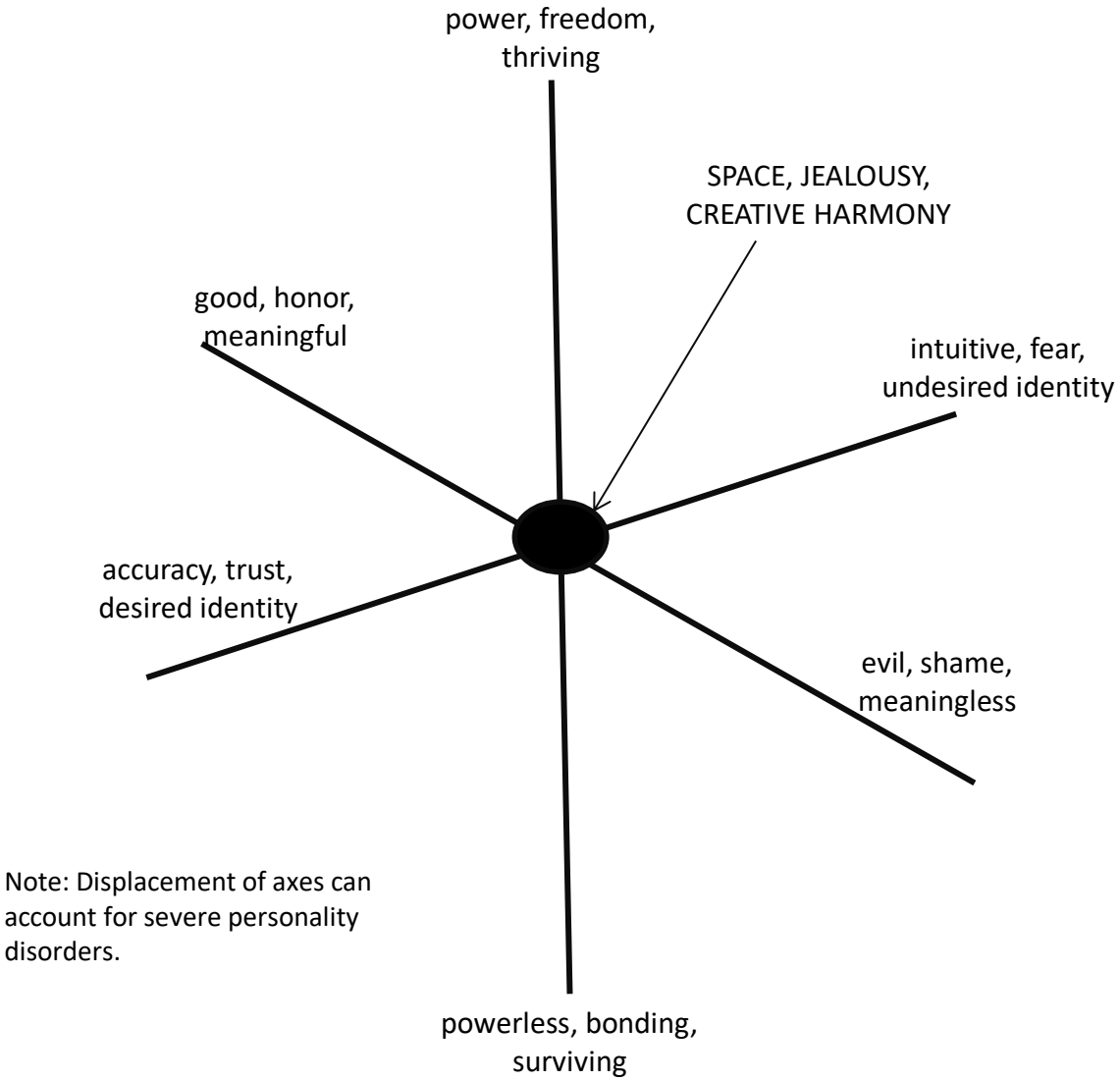


Figure 3: Overlapping Axes of the Thought Dynamo

Figure 3 depicts the intersecting dynamic of the three sets of axes. Good, honor and meaningful interact on a continuum with evil, shame and meaninglessness. Powerful, freedom and thriving interact of a continuum with powerless, bonding and survival. And accuracy, trust and desired

identity interact on a continuum with intuition, fear and undesired identity. The central construct of these intersecting axes is creative harmonious, jealous space. The logic of emotion, logic of intellect and imagined outcomes are therefore conceptualized to interact with and influence each other. The relative weights of these factors provide the difference in cultural decision making preferences.

Regarding the paradox of full, partial full-empty and null center sets

As a comparative example, it may be noted that a significant difference between core tenets of Judaism and Buddhism may be found in the central construct of the logic of emotion. A central teaching within Judaism is the Ten Commandments. The core emotional aspect of God revealed in the Ten Commandments is jealousy. "You shall not make to yourself an idol in the form of anything in heaven above or on the earth beneath or in the waters below. You shall not bow down to them or worship them; for I, the Lord your God, am a jealous God…". A core tenet of Buddhism is emptiness of all including emotions. Tarthang Tulku (1978), in his work Openness Mind, on explaining Tibetan Buddhism to a Western audience states, "The space between thoughts has a quality of openness which is very close to emptiness. This space is not caught up with discriminations or obscurations. Reaching it is like diving deep into the ocean; there is a vast stillness." The paradox of jealousy (full set) and emptiness or ambivalence (null set) can be modeled as a central construct of emotional reasoning.

The central construct of the logic of intellect can be modeled as empty, partially occupied or full space and the central construct of imagined outcomes can be modeled as creative (orderly to chaotic) harmony over time. Thus, the central constructs of the mind are jealous space seeking creative harmony.

A Case for Pain and Pleasure

A pain-pleasure continuum is a panhuman characteristic of the mental world that is an interpretation of movement of physical particles. These interpretations are subject to re-definition (e.g. the negation of pain as practiced by various religions). However, in general, pain and pleasure are viewed as sensory feelings that are pan-human while emotions (higher levels of abstraction of pain-pleasure) are viewed as culturally constructed interpretation of pain and pleasure feelings.

Where do pain and pleasure plotted on the decision making axes? Pain and pleasure are preliminary categorizations of sensory input. Sensory pain interpretations (such as stumping a toe on a rock) are generally plotted in the bonding, fear and possibly mildly shameful quadrant of emotional logic. Sensory pleasure interpretations (such as drinking a great cup of coffee) are generally plotted in the freedom, trust and mildly honoring quadrant of emotional logic. However, this feelings continuum of pain-pleasure does not form a sufficient basis for understanding the complexity of emotional decision logic. Though all input is viewed as pre-sorted onto pain-pleasure continuum by the mind, the culturally constructed categories of emotions based upon the pan-human axes of logic of emotion are posited as a more sufficient means of categorizing sensory input at the emotional level.

A Case for Beauty and Perfection, Passion and Love

Notably absent is the concept of beauty within the decision making framework. The beauty-ugliness continuum can be constructed from the central constructs of the above three interconnected axes. That is, beauty-ugliness can be seen as a spatial creation with various aspects and degrees of jealousy and harmony (at full, null or partial set perspectives). This accounts for numerous decisions that are made based on beauty (e.g. purchase of vehicles, marriage partners, vacations, foods). Perfection can be viewed as a descriptor of beauty or beauty of perfection.

Passion can be viewed as an intensity level involving beauty. And love can be viewed as a formation of a jealousy bond that includes both passion and beauty.

Holographic Neural Networks and Behavioral Extensions

The brain's ability to relearn functions after brain damage and it's ability to function after one hemisphere has been removed and the reported holographic nature of the brain – all point toward an embedded infinities modeling of the brain by the mind as an reliable description of observable inputs into the brain and interpreted by the mind.

This embedded infinities modeling of neural networks would also extend to neural nets in the spine and stomach regions of the body. These areas have been conceptualized to yield "gut" knowing by various cultures.

The mind influences the physical neural network system that in turn influences the behavior of the physical body (from arm movement to speech). Thus, the mind also reaches beyond the mind into the beyond it perceives for interpreting sensory input from an embedded infinities perspective.

Probabilistic Cause and Effect Embedded in Time

Is time an embedded infinity within the mind and the mind's conceptualization of beyond the mind? Is the present embedded in the past; the past in the present; the future in the past and the present?

The mind can perceive the future in the present. Imagined outcomes, with probabilistic cause and effect over a period of time, are the minds attempts to see into a future with consequences from the past and the present and future present effects. The embeddedness of time as perceived by the mind allows for such thinking. [Time is infinitely embedded in space and space is infinitely embedded in time from an abstract-limitations-of-the-mind perspective.]

Across cultures, time is perceived differently and with different values. However, all of those constructs are embedded within an embedded infinities model of the mind that allows for probabilistic cause and effect over varying periods of time.

Metaphysical Thought – the 'Beyond' of Embedded Infinities

A model of the mind must grapple with the human mind's compulsion to grapple with metaphysics – beyond the physical, beyond current states of consciousness, beyond death. Is there a God? Is there a spirit or soul? Is there an afterlife? Is there an eternity or a heaven or a nirvana? To varying degrees all cultures have addressed such questions with varying answers.

The question of why the mind plays in the arena of metaphysical thought may be addressed from the perspective of embedded infinities. Discrete embeddedness allows for distance – infinite distance. Infinities allow for beyond a finite distance and infinite numbers within finite boundaries. What then is perceived is a beyond. What is known regarding the beyond is that it is perceived from the mind that universally is posited to be constructed to function from an embedded infinities perspective.

Is the spirit/soul the beyond of the mind? Is there any more that can be known of a beyond? Only that any mental knowing in this life would be from an embedded infinites perspectives – the abstraction limitation of the mind. Even perceived revelations from a metaphysical world would enter human knowing through the abstraction limitation of the mind. All such knowings are subject to the consideration of reliability of knowing.

Reliability of Knowing the Mind and Beyond

The primary limitation of a model of the mind is the assumption that the mind cannot think outside itself. It does maintain external prompts, e.g. a picture may remind one of a previous event or a series of mathematical formula, but the mind doesn't think outside itself.

So what conclusively can be stated about the mind and beyond? Only through trust and with varying degrees of uncertainty can one 'know' any mental reality or a beyond the mind reality. The mind, therefore, is dependent upon its concept of trust and the sensory input from the perceived external in order to reliably know itself and a beyond reality. Observable, experiential data (within probabilistic cause and effect parameters) fortifies trust.

The mind must trust itself to some degree or insanity will ensue by degree. Likewise, trust makes deception, as judged by the community of minds weighing probabilistic cause and effect, a possibility.

Part Two: Brief Applications for a Model of the Mind

Brief applications for an embedded infinities model of the mind will include: global economic theory, market research strategies, philosophy, metaphysics, psychological health, educational agenda and ethics, social health models, intercultural communication and relations, international conflict negotiations, physics, biology and artificial intelligence programming. A brief research agenda will also be included.

Global Economic Theory

Much of capitalistic economic modeling deals with the accuracy of economic indicators including consumer confidence, i.e. the fear-trust continuum regarding purchasing. This model of the mind suggests that other decision making axes such as honor-shame and identity are also powerful forces within any economic system. Nationalistic identity and personal shame management motives may drive certain markets as much as consumer confidence.

Market Research Strategies

A model of the mind yielding 3-dimensional embedded decision making axes as previously described can form a basis for market research. Researching the decision making preferences of logic of intellect, logic of emotion and imagined outcomes of market segments can form a competitive edge within capitalistic markets and make mass customization more attainable.

Philosophy

Epistemology includes the study of reliability of knowledge – "How can I know if I know anything?" The subject has often been dominated by formulating and discussing logical rules of intellect. Such rules are a pursuit of accuracy. This model of the mind suggests that reliability of knowledge is based not only on accuracy of facts and intellectual rules of logic but also on the other continuums within the model. All language across cultures have a similar decision making basis. Wittgenstein came to the same conclusion in Philosophical Investigations, but he stopped short of producing a model for the logic of emotion and imagined outcomes that suggests a paradigm of rhetoric.

A culture-general embedded infinities model of decision making – the core process of the mind and thus the core process of culture formation – incorporates a basis for understanding decision making about knowing at a level of congruence, contradiction, paradox, comedy and disillusionment and postulates a means of viewing the rhetoric of the mind (three interconnected and interacting axes of decision making).

Postmodern philosophy supposes that all theories of knowledge are founded on an illusion that time is continuous. This re-qualification of presence from the modern perspective provides a basis for crumbling all theories of knowledge. The embedded infinities model of the mind and probabilistic cause and effect (embedded within the model) would suggest an embedded infinity of time (within the mind). This model would then account for the possibility of a theory of knowledge since embeddedness allows for connectivity and interaction of discrete infinities of time and space rather than only discrete infinities of presence that create an illusion of connectivity. This connectivity provides the possibility of knowings between people and across cultures over varying periods of time while allowing for interpretation and fuzzy reliability based substantially on the emotional construct of trust and refined by probabilistic cause and effect over varying periods of time.

Furthermore, determinism and free will may be viewed as an internal shifting from an innate rhetoric of the mind (determinism) to the implementation of that rhetoric (free will). Symbolic reasoning (based in the spatial nature of words) then links that rhetoric to a beyond the mind – the question of metaphysics.

Metaphysics

Approaching the metaphysical through a symbolic superstructure of embedded infinities suggests many questions. Does space mirror the mind or does the mind mirror the nature of space? Is space unitary in nature or triune or something else? Is symbolic reasoning of superstructures of the mind supported by probabilistic cause and effect observations or is it delusional fantasy? Does a superstructure of embedded infinities suggest a Deity, Deities or that the superstructure is a Deity or that no god exist? Has God created the universe in a way that suggests his fingerprints as a singular God or a plural gods/goddesses or a triune God? In the Christian faith, God has revealed Himself as triune (three in one) and Jesus as the consummate embodiment of embedded infinities (i.e. God-Man). Furthermore in Christianity, God has revealed himself as the Jealous, Omnipresent (in space) and the Creator of creative harmony (through time). Are embedded infinities his fingerprints from a Christian worldview perspective?

It is not the intent of this work to address each of these questions, but rather to suggest that an embedded infinities model of the mind provides a rich link to the metaphysical pursuit of beyondness.

Psychological Health Theory

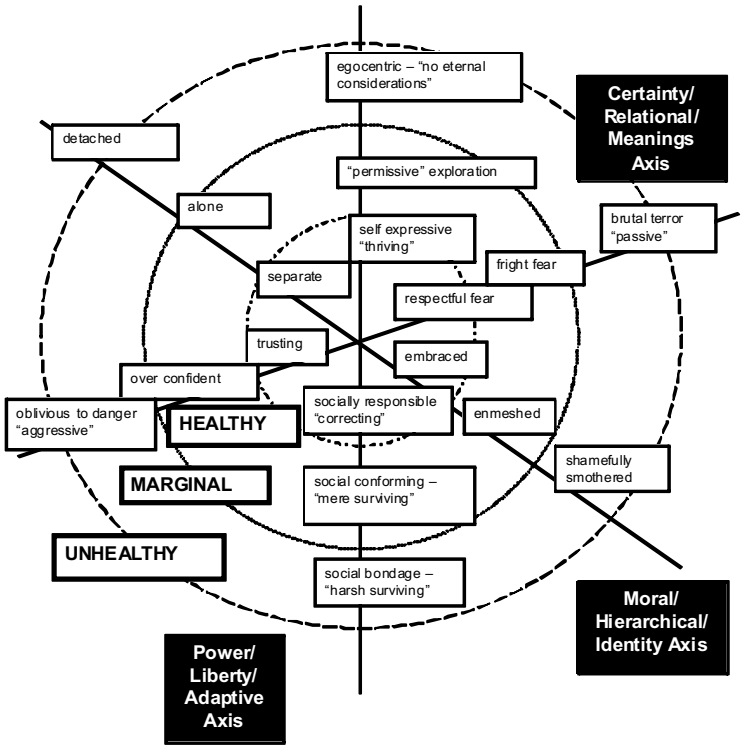


Figure 4: Three intersecting axes of psychological health

A model of the mind should speak to a psychological healthy-unhealthy continuum. Using the (see figure 10) intersecting axes of decision making (with the possibility of separation and/or rotation of axes) mental health categories can be constructed as well as mental illness categories. Also these constructs give bases for empathetic connectivity between people.

Educational Agenda and Ethics

The goals of education involve the mind. So a model of the mind has impact on an educational agenda. Learning styles can be construed from the axes of intellect, emotion and imagined outcomes. The styles can impact how knowledge is developed and delivered in the education arena.

Education recognizes underdeveloped mental capacities. Mental abilities can be conceptualized as each of the nine axes and three central constructs (e.g. moral reasoning ability, spatial reasoning ability, emotional reasoning ability). The development of each of these abilities becomes an educational agenda.

As an example of education regarding a moral reasoning ability to discern good and evil, the below 19 perspectives of ethics is posited. These perspectives are constructed from the three intersecting axes systems within the overall flow of decision making – the core process of decision making.

1. Concepts of fairness and revenge as determined by judging others' behavior toward oneself
2. Concepts of belief and symbolic meanings as determined by religious beliefs, cultural values and symbols regarding perceived reality
3. Concepts of verbal and non-verbal truth telling with consistency as determined by behavior conformity over time
4. Concept of continuums and paradoxes as suggested by relative degrees of morality and belief paradoxes
5. Concept of intent as determined by judging motivation (decision preferences)
6. Concept of protection of innocents as determined by powerlessness
7. Concept of probabilistic cause and effect as determined by experience
8. Concepts of relational loyalty and jealousy as determined by in-group relations and dynamics
9. Concepts of emotional integrity and management as determined by individual emotions values and group emotional patterns
10. Concepts of power and authority as determined by beliefs, organizational practices, relationships and change dynamics
11. Concept of goal achievement (thriving) as determined by goals and value of achievement
12. Concept of necessity (surviving) as determined by physical and psychological needs
13. Concept of ownership (physical and mental space) as determined by individual and group agreements
14. Concept of identity preservation as determined by desired and undesired identities
15. Concept of law (permission, restriction, innocence, guilt, condemnation and mercy) as determined by rules of behavior previously agreed upon by a group
16. Concepts of honor and shame as determined by a hierarchical system (of people and/or codes) established by society or religious revelation
17. Concepts of beauty and emptiness as determined by creative harmony of jealous space

18. Concept of harmony/health as determined by system equilibrium
19. Concept of relational love as determined by individuals and culture

Societal Health Theory

Formulating a model of societal health is very problematic. However, psychological health models are common. Therefore, the intersecting axes of decision making (see figure 10) may also form a model for discussing societal health. This model is presented not as a definitive description of or prescription for societal health but as an illustration that embedded infinities model of the core process of the mind (i.e. decision making) can form a perspective for discussing the topic of societal health

Intercultural Communication and Relations

The field of intercultural communication and relations is predicated on the core process of the mind – decision making. Communicative decisions that impact relationships are a fundamental aspects of the field.

An embedded infinities model of the mind with a model of decision making that is culture-general can affect the research agenda and practice of intercultural communications and relations. Cultural differences in decision making preferences are an important difference between cultures. This model of the mind accounts for decision making commonalities and allows for research on decision making differences.

International Conflict Negotiations

The model can be useful to thoughtful people as they approach difficult international conflict issues. Specifically, the imagined outcomes axes of decision making can be used to construct a model for relating and conflict resolution. These styles – based on a culture-general model of decision making – can be used to help negotiations across cultures. The styles match the eight quadrants and the central construct of imagined outcomes (see figure 11).

	Meaning	Identity	Adaptation	Style of Relating to Others
1	High Value	Low Value	Low Value	Accommodating
2	Central	Central	Central	Adapting
3	Low	Low	Low	Avoiding
4	High	High	High	Collaborative
5	Low	High	High	Competitive
6	Low	High	Low	Compromising
7	High	High	Low	Sacrificing
8	High	Low	High	Seeking
9	Low	Low	High	Surviving

Figure 5: Conflict Resolution and Relating Styles

In addition, decision making preferences of each party in conflict can be mapped and alternative decision making approaches considered. For example, honoring the powerless may be a back-door resolution approach to a conflict positioned by moral and identity management decision making that is emotionally fueled by revenge for shame inflicted.

Physics

Can a model of space that is reliable to observable data be formulated from a discrete, (3) 3-dimensional, embedded infinities model? That question will be explored in a future paper.

Biology

The study of life can be constructed as a study of embedded infinities – life reproducing itself from the life embedded within DNA. Only that which is embedded can come forth – with infinite variations. Thus, the study of animal minds and behavior can be viewed from an embedded infinities perspective.

Artificial Intelligence Programming

Any attempt to program computers with artificial intelligence, implies a model of human intelligence. The embedded infinities model of the mind and its application to decision making has relevance to AI programming. The potential of mapping the human mind (individually and/or collectively) by employing an embedded infinities model of the mind can lead to fuzzy optimization decision suggestions from AI programming.

Bibliography

Decision Making Literature

Byrnes, J. (1998). The nature and development of decision making: A self-regulation model. Mahwah, NJ: Lawrence Erlbaum Associates.

Maruyama, M. (1987). Communication between mindscape types. In J. VanGigch, (Ed.), Decision making about decision making: Metamodels and metasystems (pp. 79-97). Cambridge, MA: Abacus Press.

Economics and Consumer Behavior Literature

Onkvisit, S. & Shaw, J. J. (1994). Consumer behavior: Strategy and analysis. New York: Macmillan College Publishing Company.

Zaltman, G. (1997). Rethinking market research: Putting people back in. Journal of Marketing Research, 34, 424-437.

Linguistics and Philosophy Literature

Derrida, J. (1973). Speech and phenomena and other essays on Husserl's theory of signs. Evanston, IL: Northwestern University Press.

Lakeoff, G. & Johnson, M. (1980), Metaphors we live by. Chicago, IL: University of Chicago Press.

Wittgenstein, L. (1958). Philosophical investigations (2nd ed.). (G. E. M. Anscombe, Trans.). Malden, MA: Blackwell Publishers. (Original work published 1953)

Mathematics

Lakoff, G. & Nunez, R. (2000). Where mathematics comes from. New York, NY: Basic Books,

Psychological, Cultural and Intercultural Communication Literature

Bennett, M. (Ed.). (1998). Basic concepts of intercultural communication: Selected readings. Yarmouth, Maine: Intercultural Press.

Clark, A. (1997). Being there: Putting brain, body, and world together again. Cambridge, MA: MIT Press.

Clanton, G. (1998). A sociology of jealousy. In G. Clanton & L. G. Smith (Eds.). Jealousy (3rd ed.) (pp. 297-312). New York, NY: University Press of America.

Ennis (2002). A theoretical model for research in intercultural decision making.

Fessler, D. (n.d.). Toward an understanding of the universality of second-Order emotions. In A. Hinton (Ed.), Biocultural Approaches to the Emotions. Cambridge, England: Cambridge University Press (forthcoming).

Levy, R. (1984). Emotion, knowing, and culture. In R. A. Shweder & R. A. LeVine (Eds.), Culture theory: Essays on mind, self, and emotion (pp. 214-237). Cambridge, England: Cambridge University Press.

Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. Genetic, Social, and General Psychology Monographs, 12 (3), 339-361.

Osgood, C. E., May, W. H. & Miron, M. S. (1975). Cross-cultural universals of affective meaning. Urbana, IL: University of Illinois Press.

Strauss, C. & Quinn, N. (1997). A cognitive theory of cultural meaning. Cambridge, England: Cambridge University Press.

Wierzbicka, A. (1992). Semantics, culture, and cognition: Universal human concepts in culture-specific configurations. Oxford, England: Oxford University Press.

Religious Literature

Tarthang Tulku (1978). Openness mind. Berkeley, CA: Dharma Publishing.

The Bible

Science and Physics Literature

Misner, C., Thorne, K., & Wheeler, J. (1997). Gravitation. New York, NY: W. H. Freeman and Company.

Gribbin, J. (1995). Schrödinger's kittens and the search for reality: Solving the quantum mysteries. Boston, MA: Little, Brown and Company.

Wolfram, Stephen (2002). A new kind of science. Wolfram Media, Incorporated.

April 26, 2008
Updated from September 5, 2003

Assumptions and Pathway for Achieving Artificial General Intelligence

Ralph C. Ennis
Raleigh, NC
ralphennis@gmail.com

DRAFT WORK – April 2013

Abstract. The purpose of this paper is to posit core assumptions for human-level intelligence (HLI) and an assumptive pathway to achieve artificial general intelligence (AGI). These assumptions and pathway can form an evaluative means for constructing decision models and algorithms for AGI.

The core assumptions for HLI revolve around three constructs: 1) the nature of language, 2) sensibility of reason, and 3) ethical reasoning. Implicitly these constructs should be adaptive to cross languages, cultures and subject matter of the human experience.

The assumptive semi-sequential pathway to AGI involves: 1) process inputs, 2) learning, 3) evaluation, 4) imagination, 5) optimization, 6) resolving conflict, 7) solidifying rules of thumb, 8) selective memory recall, 9) prediction, 10) explanation.

From these core assumptions a mean of describing process (thought) decisions and output (behavior) decisions can be forthcoming. The assumptive pathway can guide formation of decision mapping structures and logical employments within those structures.

Keywords. Assumptions, pathways, artificial general intelligence.

1 Introduction

In 1950 Alan Turing established a test to evaluate the quest for human-level intelligence by machines or artificial general intelligence (AGI). That test involved a blind dialogue between an examiner, a human and a machine. When the examiner can discern no difference within the dialogue between the human, the machine and himself, then human-level artificial intelligence has been achieved. This simple test of "can't tell the difference" has been a benchmark for achieving AGI. Over the past six plus decades since Turing proposed this test much progress has been made and yet this goal remains elusive.

Many assumptions are embedded within the Turing evaluative procedure. The foremost assumption involves determining the bare essentials for human-level intelligence. After which an assumptive pathway is required to achieve such intelligence before algorithms and code can be written. In most human goal-seek ventures, if the assumptions for solving a problem are both comprehensive and effective, the likelihood of achieving the goal is greater than with less effective or comprehensive assumptions.

The subject of this paper is the core assumptions for human-level intelligence and an assumptive pathway to achieving AGI. And the primal assumption is that if the core assumptions and the assumptive pathway are comprehensive and effective, then the probability of achieving AGI increases.

2 Core Assumptions for Human Level Intelligence

Three core assumptions of human-level intelligence (HLI) will be discussed: 1) the nature of language, 2) sensibility of reason, and 3) ethical reasoning. Stated otherwise, HLI is associated with language that makes sense and inherently involves ethical reasoning. Non-human intelligence may operate without this full set of assumptions. Though a different list could be devised, this set is posited as fundamental to human intelligence and should be included in any discussion of AGI.

2.1 The Nature of Language

The nature of language is complex across cultures, age, demographics and non-verbal expressions. Since Turing's test requires dialogue, for this discussion words will be viewed as the primary component of language for any AGI.

Language, navigated through words, is symbolic, spatial-temporal, contextual and requires authorship. A "word" primarily represents, not itself, but something distant, apart from itself. That word is symbolic—it represents some spatial-temporal construct at a concrete level (often referred to as a sign) or a more abstract connotations (symbol).

> The philosopher Wittgenstein (1958) referred to the "spatial and temporal phenomenon of language" (p. 47). This spatial-temporal quality of language allows an author of words to transcend himself—perceive and transmit beyond himself. [Various spatial-temporal vantage points thus accounts for the differentiation between self and other awareness necessary for any discussion of commonly shared (perceive) reality as well as a sensibility for a variety of spiritualities].

A word is also contextual. The same word in two diverse contexts may carry different, though maybe similar, meanings. Furthermore, the diminishing presence from authorship to the receiving party accounts for much confusion in the transmission of meanings in language between humans.

For example, the simple words "I love you" conveys deep meanings to most who hear it or long to hear it. These words convey difference nuance meanings when spoken to a loving spouse as compared to a beloved child. Each word in this sentence communicates a spatial-temporal construct. "I" and "you" are concrete symbols of distinct persons while "love" is an abstract symbolic construct that has found meaning over time (to the author and hearers) through acts of love that resonate sensibility to them. Furthermore, inherent within these words is a sense of the author's presence in time. If "I love you" is spoken and received in the immediate present, the meanings is full and rich to the receiver. If however, it is spoken and received a year later without any direct author presence, the meaning of "I love you" may be quite different. If the author has subsequently abandoned the intended receiver, the receiver will undoubtedly understand these words quite differently.

Thus, if AGI code can't convincingly convey words originating from itself and commonly understood symbolically, spatial-temporally, and in context, then Turing's test may not fully satisfied.

Language is embedded within diverse emotional constructs across cultures. Every healthy human being experiences emotions. Nevertheless there is no uniformity of emotional words that apply across all cultures.

For instance, in the Japanese construct of emotionality "*amae*" is a powerful emotion. The Japanese psychiatrist Takeo Doi (1981) unpacks this emotion for Westerners by stating, "The Japanese term *amae* refers, initially, to the feelings that all normal infants at the breast harbor toward the mother—dependence, the desire to be passively loved, the unwillingness to be separated from the warm mother-child circle and cast into a world of objective 'reality' " (p. 7). He went on to say, "… all the many Japanese words dealing with human relations reflect some aspect of the *amae* mentality. This

does not mean, of course, that the average man is clearly aware of *amae* …" (p. 33). In the English language there is no direct translation for *amae*.

The complexity of language is displayed in the fact that all language has an emotional component in the originating and receiving for words—and that emotionality must be accounted for in AGI. Sometimes that emotionality may seem to be non-existent but it is better perceived as muted or laying in wait to spring into action—nonetheless, emotions are always present in language. Stated another way, the regions of the brain responsible for emotional processing never go entirely dormant. [Even mathematical symbols must pass through the grid of a students' emotionality as they struggle to solve problems.]

Thus, Turing implicitly requires that the examiner notice an appropriate handling of the emotional world between the dialogue of human and machine. For full AGI to be achieved the code must account for a multiplicity of nuanced emotions across cultural contexts.

Human language is bodily encased. Language is experienced and transmitted in and through the body that innately perceives "attractions-aversions." Without logical awareness, a child responses to stimuli in a manner that resembles the reaction of simplest of life forms to outside influences. Each move toward or way from stimuli it innately perceives as beneficial or threatening. This surviving-to-thriving reaction is often translated into a language of "pleasure" and "pain" or, at higher abstract levels, into "attractions" and "aversions" or "harmony" and "dissonance."

Achieving AGI does not necessarily require code embodied by sensory "flesh." Nevertheless, AGI must account for "pain and pleasure" at a primal level. Without such primal responses, Turing's examiner may eventually perceive a flaw in the machine that doesn't account for HLI encased in bodies that perceive beneficial and threatening stimuli.

Words are best processed for learning as they attach to images and meanings attached through analogy. People process and remember images far better than words (Grady, 1997). The symbols that words reflect are often birthed through images. Images are powerful. They drive much human communication and learning. By comparing images, humans use pattern recognition to associate words, form abstractions, and learn through analogy. Gentner, et al. (2006) states, "The proposal that comparison processes can promote language learning is based on research in analogy and similarity." And Marvin Minksy (2007) believes only through the pathway of analogy will AGI be achieved.

For instance, in the mind of a child, the image of mother's face becomes deeply associated with the word "mommy." In time, the pattern of woman-with-baby recognized in living images (or artistic displays) is generalized to the abstract of "Mother." And by analogy "Mother" can extend to any female animal with offspring or even to Mother Earth as a birther and nourisher of life.

The fluidity of learning required for AGI may best be negotiated through image associations and secondarily through word associations until analogies are formulated that facilitate abstractions. The Turing examiner will look for such learning abilities within the dialogue between human and machine.

Words are not discrete and no discrete definition of terms is required as a starting point for AGI—rather a process for dynamic adjustment of words is required. As Jacque Derrida, the late French postmodern philosopher, has stated, "It is at the price of this war of language against itself that the sense and question of its origin will be thinkable … Language preserves the difference that preserves language."

This convolution of words is what the twentieth century philosopher Edwin Wittgenstein referred to as "Language is a labyrinth of paths. You approach from one side and know your way about; you approach the same place from another side and no longer know your way about "(Philosophical Investigations 203). Or one might say, language is not formed by discrete, immovable categories

(words) but rather by flows of embraced constructs through continuums of intersecting and interacting pathways.

Fixed definitions of words will insure that AGI will not be achieved. However, a radical relativism approach to language also dooms the quest for AGI. Common sense (semi-ambiguous) meanings inherent in a relational usage of words is a better approach both for transmission of meanings and for learning—and will appease the Turing examiner.

2.2 Sensibility of Reason

In reasoning with a child (or across cultures) is becomes quickly apparent that sensibility is fluid. Or stated another way, two people can arrive at the same or different conclusion by very different pathways—and yet sensibility is achieved for both people.

The discussion on sensibility usually starts with rules of formal logic. Computer code usually starts and stops there. However, human-level intelligence is present long before formal rules are acquired or followed. What makes sense to one three year-old may not make sense to another three year-old and yet both are seeking "sensibility"—trying to make sense of their worlds.

One possibility is that sensibility is the play of dissonance and harmony with the energy frequencies within the brain. Minsky (1981) has suggested a link between music and meanings. Recent work by Lu, J. et al. (2012) has translated brain waves to music, this avenue to sensibility must be explored. If fruitful, we might view the brain as a "music box" continually seeking harmony while resolving even-present dissonance. This play of resonance may account for sensibility, irrationality and even disease.

In any case, Turing's examiner would surely ask both human and machine the question, asked with annoying frequencies by most three year-olds, "why" and expect something "sensible" response in the dialogue.

2.3 Ethical Reasoning

In a previous paper (Ennis, 2013) I noted:

> *Mikhail (2007) frames this poignant question relevant in the pursuit of an ethically-based artificial general intelligence (AGI): "Is there a universal moral grammar and, if so, what are its properties?" Stated otherwise, is there a set of rules that govern the formation of all ethically acceptable behaviors across cultures?*
>
> *Evidence can be found on any kindergarten playground across the global community that ethical reasoning is at play. In what part of the human experience is some construct of "fairness and harmony" non-existent? This construct may seem suspended or violated at various times, but an innate awareness of fairness and harmony resides within us all—even in our early childhood interactions (Smith, et al., 2013).*
>
> *Fairness may be defined differently across individuals, families and cultures, but yet it resonates within all social structures even if pathways to it are blocked. Fairness to some implies non-bias equality of quantity and quality. However, this definition rarely works out well without the consideration of context.*
>
> *For instance, is it fair to an eight year-old sister to be treated equally with her four year-old brother, or vice-a-versa? Most parents would conclude unequal treatment is far more "fair" that an unwavering pursuit of equality. Much to the consternation of young siblings, most parents conclude that it does not have to be equal to be fair. Fairness is contextual to age, abilities, available resources, etc.*

If fairness is not somehow achieved or at least approximated, we humans recognize that harmony (dynamic balance) within a system may be threatened or disrupted. Back to the family system—sibling disputes over fairness can disrupt the sense of harmony for all in the family.

What remains in the pursuit of ethical reasoning is not the question of a set of ethical rules that are proven to be universal, but rather can a grammar, a DNA, for ethical reasoning be established from which a diversity of contextual rules might be fashioned and situations evaluated for ethical acceptability? Is that DNA applicable in the formation of ethical rules and parsing of existing rules across cultures—even when the rules seem in conflict?

A solution to that ethical DNA and subsequent management of it is paramount in the quest for AGI that accounts for the human sense of fairness and harmony and the hope that it becomes ethically equal or superior to humanity. As Pana (2006) states, "We do not have to implement a moral code, but to create a moral intelligence, we can aspire to a condition of potentiality, not the generation of some fixed reality."

The examiner of AGI will quickly perceive the ability of the human to seek fairness and harmony. But will the machine pass this test? The answer is or should be of upmost concern for all in the enterprise of building AGI systems. Without ethical reasoning, AGI may be very intelligent but it will not resemble child or adult human-level ethical intelligence. Such intelligence may find no difficulty in prescribing and enacting decisions that humanity may find utterly unethical and disastrous.

3 Assumptive Pathways to Artificial General Intelligence

With the above assumptions in mind, a pathway to AGI is suggested below.

3.1 Process Sensory Input

All forms of sensory input (visual, auditory, taste, touch, smell) must eventually fit within a model for AGI. However, visual images and written words seem sufficient to begin. These inputs must be received from those inputting data or auto-gathered across data fields. The inputs must then be discarded as appropriate or filed in retrievable though adjustable filters.

3.2 Learning

Learning occurs through processing new input. It also occurs through processing feedback from previous decision consequences. These inputs allow for reinterpretation of prior inputs and formation of rules of thumb—held as solid but adjustable.

Learning can occur through analogy. This comparative process might best be negotiated by parsing words and images using an ethical DNA and by using spatial-temporal recognition patterns.

The grammar rules for word order in sentence context are learned rather than assumed or pre-programmed. And in like fashion, formal rules of logic are learned.

Auto-learning can occur through auto-gathering and processing of data. This data is temporarily mapped and adjusted into a more appropriate location by rules of thumb related to the words or images in context.

3.3 Evaluating Decisions for Ethical Acceptability

As much as ethical reasoning has fallen out of favor in our current post-modern rationality, even the construct of tolerance is heavily laden with ethical acceptability. Some means of evaluating process and output decisions for ethical acceptability must be achieved.

3.4 Imagination of Possibilities

As inputs increase and are linked within varying emotional intensities, imagination (i.e. dreaming) becomes possible. Even as a three-year old child lives in an imaginative play world, so AGI must have an ability to imagine what is not actual. Without imaginative powers, AGI will fail in the eyes of the Turing examiner.

3.5 Optimization of Decisions

Optimization of decisions is a truly human intelligent pursuit. Achieving a goal involves uncertainty. People seek the best result. That best involves both sound reasoning that is congruent with prior rules of thumb and mindsets as well as, in time, a positive evaluation of decision outcomes. Congruence decisions can be conceptualized as acceptable.

Dissonant decisions are conceptualized from warning to dangerous. Dissonance might possibly lead to a redistribution of rules of thumb. Comic decisions are conceptualized to facilitate dissonance and redistribution of rules of thumb. [The use of human comedy to create cognitive dissonance is the art form of comedy shows.]

Paradoxical solutions are conceptualized as optimal. The adult human mind runs headlong into paradoxical reasoning. That is, thoughts A and B are held to be congruent when viewed separately but when viewed together they seem contradictory (dissonant). Often the optimal solution for a system may appear paradoxical. A paradoxical conclusion can be seen as a means of declaring the limits of the human mind to solve a problem that is based within our spatial and temporal limitations. Optimizing decisions in global economics and physics can be viewed through this paradoxical pathway.

Turing's examiner may well pass AGI that resembles child-like thinking because paradox is seldom on a three year-olds mind. However, true adult-level AI must account for paradoxical optimization—the best solution is sometimes a paradoxical conclusion.

3.6 Resolve Conflict

Every human being experiences the quandary of arguing or not agreeing with self. For instance, few people hold the same rules of thumb for social interaction at age 20 that they held so tightly at age ten. Added to this conflict are conflicts with other mindsets (individuals, nations, etc.). Human intelligence gives considerable energy to resolving conflicts within and between mindsets. Lack of resolution can have mild to disastrous results from individual mental confusion to wars between nations.

Conflict resolution within and between mindsets may possibly be negotiated through the weighted influence of mindsets and a paradoxical central construct of ethical acceptability in order to diminish dissonant conflict. AGI must pass this test as well.

3.7 Solidifying Rules of Thumb

In order to make fast decisions, human being establish rules of thumb rather than sorting through all data inputs to re-logic every decision variation. We all have our rules of thumb for what behavior to employ when it is raining. And we usually defer to those rules of thumb rather that process all available data regarding water composition, rate and velocity of rainfall, etc. before making our clothing choices on a rainy day.

Types of rules of thumb can be conceptualized to include: innate (ethical reasoning DNA), metaphoric (simple comparative rules), situational (simple consequential rules) and abstract (complex comparative and consequential rules). Abstractions, refined through sensible analogies, facilitate formation and adjustment of rules of thumb.

AGI may best be built by forming rules of thumb from a baby-mind to adult-level intelligence (vs. dumbing down from adult to child). This fragile process must be overseen and adjusted as AGI accounts for pain-pleasure in the human experience.

Any examiner will perceive the use of rules of thumb. Omission of rules might be detected by the onslaught of data a machine might use to justify their thought and output decisions.

3.8 Selective Memory Recall

Recall in humans is always selective; commonly referred to as selective memory. Though some human brains have been shown to have total memory across a progression of time, few humans actual possess such total recall ability.

AGI should then be able to recall context and time-appropriate information. A dialogue with an AGI machine will reveal an ability or lack thereof to recall this type of information and then associate it with appropriate rules of thumbs.

3.9 Prediction of Process and Outcome Decisions

The foundation of predicting decisions is probabilistic cause and effect of imaged decision outcomes adjusted through the feedback of prior decision consequences. Within this rubric prior decisions are factored in but less predictive. Thus, parents might predict (with some degree of probability) that their three year-old son will decide to eat all of his vegetable today because he responded so well to negative consequences from last night's traumatic dinner experience at Grandma's house.

AGI will demonstrate some ability to predict. Whether successful or not, the propensity to predict is inherent in human-level intelligence.

3.10 Explanation of Decisions

AGI must have sensible reasons to some discernable degree. Decisions without an articulated rationale is less that human-level intelligence. These explanations can be conceptualized as congruent or dissonant with prior data associations. The play of congruence-dissonance might be negotiated through an approximation of music from brain waves.

4 Conclusion

The implied goal of this paper is that an AGI software program incorporating the above core assumptions and assumptive pathway will achieve human adult-level artificial intelligence after much input has been processed, the ethical DNA has been established and many situational and abstract rules of thumb have been formed and refined.

Future research can employ these core assumptions as a mean of describing process (thought) decisions and output (behavior) decisions. And the assumptive pathway can guide formation of

decision mapping structures and logical employments within those structures. [A precursor to a mapping model is put forth by Ennis, 2004.] These structures are best formulated to interface with future mapping of the neurons of brain and possibly employing memristors as a hardware means for better storage and manipulation of data that is often described on continuums rather than discretely.

Other assumptions and pathways may indeed be needed to fill in a road map to AGI. The assumptions put forth in this paper, I maintain, are essential to passing the Turing examination.

In addition to Turing's test, true HLI must account for irrationality and unethical behavior while hopefully presenting a means of moderating such human tendencies. Within the above assumptions for HLI and the assumptive pathways to AGI both tests are within view, even if not immediately achievable.

To that amazing goal, the field of AI continues with an uncertain end regarding success and the desirability of that achievement. May AGI achieve not only adult-level human intelligence but also the ability to perpetually seek paradoxical ethical optimization in ways that support fairness and harmony between human and machine desire for survival and thriving.

References

Derrida, J.: *Speech and phenomena and other essays on Husserl's theory of signs.* Trans. David B. Allison. Evanston, IL: Northwestern University Press, (1973) 14.

Doi, T.: *The anatomy of dependency: The key analysis of Japanese behavior*. Tokyo, Japan: Kodansha International (1981).

Ennis, R.: A theoretical model for research in intercultural decision making. *Intercultural Communication Studies.* 8 (2004) 113-124.

Ennis, R.: An ethical DNA model for artificial general intelligence. (2013 unpublished paper)

Gentner, D. and Namy, L.: Analogical process in language learning. *Association for Psychological Science*, 16(6) (2006) 297-301

Grady, C., et al.: Neural correlates of the episodic encoding of pictures and words. In *Proceedings of the National Academy of Sciences of the USA.* 95:2703-2708. (March 1998)

Lu, J. et al.: Scale-free brain-wave music from simultaneously EEG and fMRI recordings. *Plos One*. (November 14, 2012)

Mikhail, J.: Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences.* 11(4) (2007)143-152.

Minsky, M.: Interview in *Discovery Magazine*. (January 2007)

Minsky, M.: Music, Mind and Meaning. *Computer Music Journal*, Vol. 5, No. 3, Fall 1981

Turing, A.: Computing machinery and intelligence. *Mind.* 59:236. (1950)

Wittgenstein, L.: *Philosophical Investigations*. Trans. G. E. M. Anscombe. Oxford: Blackwell Publishers. (1999)

Ethical DNA Model for Artificial General Intelligence

Ralph C. Ennis
LEAD Consulting, 9418 Harvest Acres Court, Raleigh, NC 27617
ralphennis@gmail.com

Abstract. An effort to understand ethical reasoning must not focus on a list of ethical rules but the underlying grammar, an ethical DNA, for the development of all ethical precepts. The purpose of this paper is to put forth a framework for ethical DNA (e-DNA) in a manner that is applicable to the pursuit of artificial general intelligence (AGI). This e-DNA model revolves around nine continuums and their intersections and interactions. The generality of any ethical DNA model is suggested only as it shows utility across cultural diversity. With the use of this e-DNA, the Japanese construct of *amae* is parsed. *Amae* is a complex construct within the Japanese society that impacts human relations—and thus ethical behavior among relations. The utility of the e-DNA model for artificial intelligence is evident in the geometric interactions between the continuums that provide a way forward in programming.

Keywords. Ethics, moral, artificial general intelligence, decision model

1 Introduction

Mikhail (2007) frames the following poignant question relevant in the pursuit of an ethically-based artificial general intelligence (AGI): "Is there a universal moral grammar and, if so, what are its properties?" Stated otherwise, is there a set of rules that govern the formation of all ethically acceptable behaviors across cultures?

Evidence can be found on any kindergarten playground across the global community that ethical reasoning is at play. In what part of the human experience is some construct of "fairness and harmony" non-existent? This construct may seem suspended or violated at various times, but an innate awareness of fairness and harmony resides within us all—even in our early childhood interactions (Smith, et al., 2013).

Fairness may be defined differently across individuals, families and cultures, but yet it resonates within all social structures even if pathways to it are blocked. Fairness to some implies non-bias equality of quantity and quality. However, this definition rarely works out well without the consideration of context.

For instance, is it fair to an eight year-old sister to be treated equally with her four year-old brother, or vice-a-versa? Most parents would conclude unequal treatment is far more "fair" that an unwavering pursuit of equality. Much to the consternation of young siblings, most parents conclude that it does not have to be equal to be fair. Fairness is contextual to age, abilities, available resources, etc.

If fairness is not somehow achieved or at least approximated, we humans recognize that harmony (dynamic balance) within a system may be threatened or disrupted. Back to the family system—sibling disputes over fairness can disrupt the sense of harmony for all in the family.

What remains in the pursuit of ethical reasoning is not the question of a set of ethical rules that are proven to be universal, but rather can a grammar—a functional ethical DNA be established? By using that DNA of ethical reasoning, can a diversity of contextual rules be fashioned and situations

evaluated for ethical acceptability? Is that DNA applicable in the formation of ethical rules and parsing of existing rules across cultures—even when the rules seem in conflict?

A solution to that ethical DNA (e-DNA) and subsequent management of it is paramount in the quest for artificial general intelligence (AGI) (Gubrud, 1997). This e-DNA should account for the human sense of fairness and harmony across a multitude of contexts. Asimov (1950) proposed such a moral code with his three laws of robotics, but we need a more fundamental code from which these laws and others might be derived. As Pana (2006) states, "We do not have to implement a moral code, but to create a moral intelligence, we can aspire to a condition of potentiality, not the generation of some fixed reality."

In this paper, I will posit an e-DNA model that has applicability across cultures and is adaptable to AGI. This e-DNA will account for human ethical reasoning and allow for such reasoning at a machine level of intelligence.

In short, the e-DNA code involves nine continuums subdivided as logic of intellect, logic of emotion and imagined outcomes. These nine continuums are considered in this paper along with three central constructs that arise from their intersections and interactions. These continuums allow for gradation of each endpoint on a linear scale. Furthermore, the logic of intellect, logic of emotions and imagined outcomes axes are non-hierarchal. All are conceptualized with equal weight in the decision making process.

2 Continuums and Central Constructs for e-DNA

The twentieth century European philosopher Edwin Wittgenstein stated: "Language is a labyrinth of paths. You approach from one side and know your way about; you approach the same place from another side and no longer know your way about "(Philosophical Investigations 203). With this labyrinth in mind, the e-DNA model is established "on continuums" rather than separate factors.

Though this approach is debatable, much of ethical reasoning fails to fit neatly within discrete categories. Humanity devises detailed laws to fulfill that sense of discrete ethical boundaries. However, even then the need for the "spirit of the law" to triumph the "letter of the law" becomes situationally mandatory in order to prioritize laws. For instance, the letter of highway laws may state a certain speed limit. However, if one needs to go a little faster to secure the life of a person with a medical emergency and without jeopardizing the life of other drivers, then most would conclude that some bending of the letter of the law (speed limit) to preserve the spirit of the law (preservation of life) is ethically sound reasoning.

The language of e-DNA will be put forth in English. However, each of the nine continuums can be translated into most languages with some degree of accuracy. This language difference must be accounted for—but not at this point. The nine continuums are grouped in three broader categories (see Figure 1): logic of intellect, logic of emotion and imagined outcomes. Each line in three-dimensional Euclidean space represents a continuum. Logic of intellect refers to the common language notion of "thinking a matter through to a conclusion without emotional bias". Logic of emotions comes into play when feelings, molded by cultural interpretations into emotional constructs, impact the logic of intellect. And finally, imaginations of probabilistic outcomes impact and adjust our intellect and emotional logics. The arrows in Figure 1 point to the intersection of three continuum which form a central construct for the logics and imagined outcomes.

For example, a society may disqualify a Judge from trying a suspected murderer of the son of that Judge. There is a high probability that the emotions of the Judge will blind him from conducting due process of law driven by a logic of intellect. Furthermore, the imagined outcome of such a trial will not serve the cause of justice among members in a society.

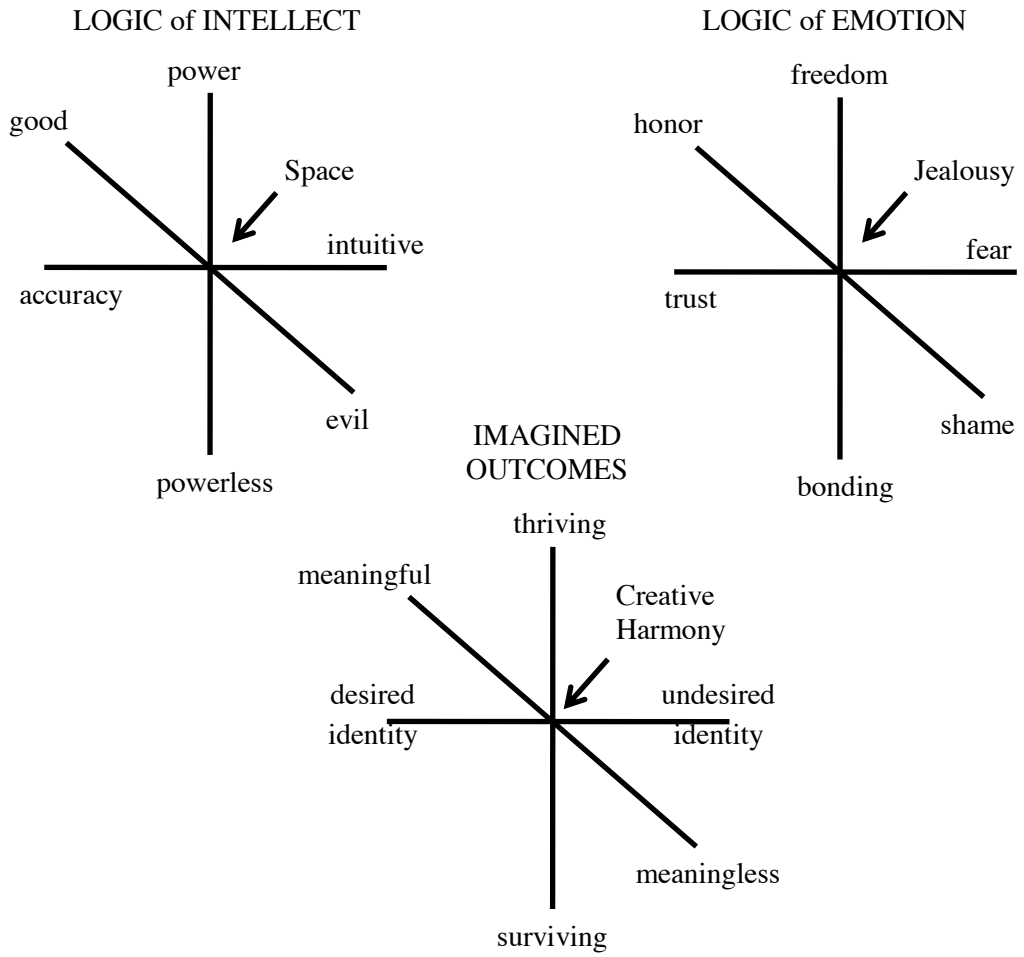


Figure 1: Ethical DNA Continuums

2.1 Ethical Logic of Intellect

Good-Evil. The very fact that all cultures have some sense of good and evil, even though they may disagree on the details, is an indication of an e-DNA. Thus a good-evil continuum simply states the obvious—humans think about ethics. This continuum must be included even if it seems redundant. However, it isn't enough to say that humans logics about good and evil. More goes into ethical reasoning than a final assessment of good and evil.

Accuracy-Intuitive. Morality includes of verbal and non-verbal truth telling that is accurate to facts or intuitively consistency with the facts. Many courts of law require witnesses to vow to tell the truth. This sense of accuracy from one's point of view is fundamental to ethical reasoning even as multiple points of view better shape an accurate account of a situation.

Powerful-Powerless. Care of the powerless, e.g. young children, is foundational to the continuation of any society. Moreover, the Hippocratic oath in medicine reasons along these lines of power. Its core tenant is to do good to patients and never harm them. This "good" is reasoned as prescribing procedures and substances to bring about better health. Better health is wrapped in the concept of power and doing harm implicitly decreases this power on a continuum to death (i.e. total powerlessness).

Space. Mental or physical spatial ownership (individual and/or corporate) is the central construct of logic of intellect. Thus space can be conceptualized as "good," "evil," "accurate," intuitive," "powerful," and "powerless." Many wars (an ethically entangled pursuit) have and continue to be fought over some conflict of space.

2.2 Ethical Logic of Emotion

Freedom-Bonding. As ethical reasoning, the continuum of freedom-bonding is best understood at the extremes of abandonment and bondage. For a parent to be totally free without any attachments is viewed by society as abandoning their child to others or to society. To be in bondage suggest varying degrees and forms of slavery. However, healthy bonding and various levels of freedoms are necessary for individuals and societies.

Honor-Shame. The management of moral behavior often comes through positive rewards that honor people or negative consequences that shame them. Sometimes the concept of authority is embedded with honor and shame. Shaming is a common form of reforming deviate behavior at home, in the classroom as well as in the broader society. Thus shame remains as an endpoint of this continuum that is the hoped for (by authorities) emotional consequence of unethical behavior. The feeling of guilt is often linked to shame. Guilt indicates lapses in behavior; shame indicates remorse in one's identity (Lewis, 1995).

Trust-Fear. A breech of trust is often considered an ethical failure. Legal contracts are formed to fortify and ensure verbal trust. Fear of the consequences of broken trust often helps negotiate trust relationship.

Jealousy. Jealousy is posited as the central construct of the logic of emotion. Jealousy has two sides – jealous *for* and jealous *of*. The latter is better referred to as envy (Clanton, 1998). To cease to be jealous *for* someone that relies on that jealousy for their protection can constitute a breach of ethics. For example, marriage is a relationship fraught with jealousy—preferable jealousy *for*, not jealousy *of*. At its best, jealousy *for* involves an emotional bonding that brings freedom, a sense of honor between members and a trust that exist when members are present or apart. At its worst, jealousy *of* can divide and destroy relationships. Furthermore, jealousy is seen to be ubiquitous in human cultures by Johnson and Price-Williams (1996).

2.3 Imagined Outcomes

Desired-Undesired Identity. To violate a person's identity through some abuse often causes strong negative reactions. Human identity structures are many and far reaching on their impact of ethical behavior. Wars have been fought to protect or advance national identities. Family inheritance laws fortify family identities within society. Certain professional identities improve the probability of securing research grants. And the imagined outcomes of present actions impact one's desired identity while decreasing the chance of an undesired identity.

Thriving-Surviving. The ethics of thriving hopefully does not value the elimination of others' survival. Humanity seeks to survive and from that basis thrive. The construct of thriving is highly imaginative. For instance, thriving in one cultural context may be imagined as possessing a cow or a bicycle. In another culture, those possessions might represent a subsistence survival.

Meaningful-Meaningless. Philosophy, art, religions are manifestation of humanity's quest for meanings that transcend themselves. Humanity, for the most part, imagines itself to be meaningful. Meaningless is conceptualized as a disruptor of productive living (thus interfering with the pursuit of thriving). Belief and aesthetic systems are designed to bring meaning into the human experience

from conception to death. To violate these meanings can be considered an immoral act. Wars have been and continue to be fought over meanings, especially religious and political meanings.

Creative Harmony. The central construct of imagined outcomes is creative harmony. This ethical concept helps maintain the goodness of perpetrating harmonious health in individuals, enterprises and societies. The violation of creative harmony—destructive dissonance—can be viewed as morally wrong under certain but not all circumstances. Civil disobedience usually seeks a better long-term creative harmony in society through a short-term pathway of destructive dissonance to reshape the rules of society. Further explanations of these continuums are put forth by Ennis (2004).

3 Central Constructs of the Continuums

The uncommon word set "creative harmony of jealous space" is achieved by overlapping the central constructs of logic of intellect, logic of emotion and imagined outcomes (see Figure 2). Ethical reasoning implies each of these ideas. Jealous space allows for property rights; both physical and mental space is inherent in the spatial-temporal nature of language. Without jealousy a sense of possession and ownership, that pervades ethical reasoning, would be a mute issue. Thus we return to the ideas of "fairness" and "harmony" in systems. The negotiation of jealousy across spatial constructs will account for "fairness" and "fairness" is mediated through "harmony" that is dynamic and thus creating new states of being across time and space.

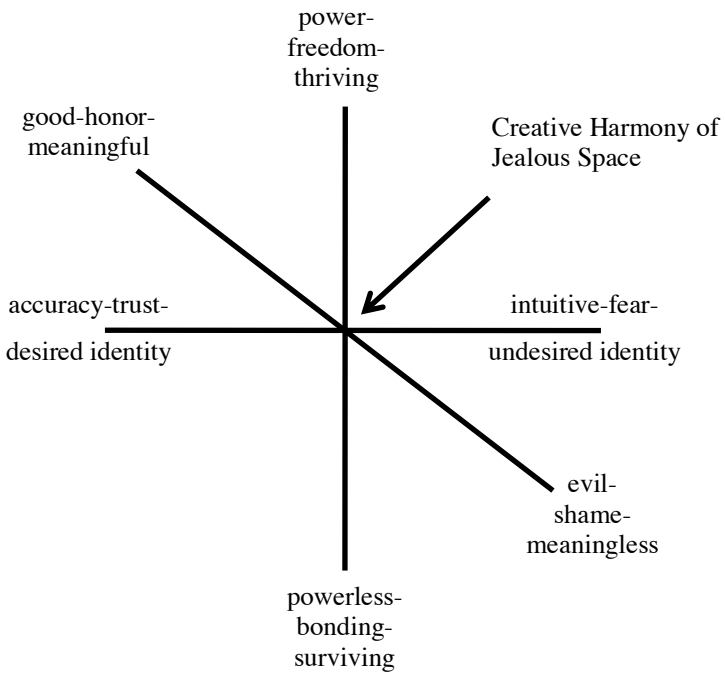


Figure 2: Overlapping Ethical DNA Continuums

4 Evaluating the Utility of e-DNA through Beauty and Love

In an effort to move from uncommon language to ordinary language, a discussion of ethical reasoning from the lens of "beauty" and "love" is needed. The previously discussed words "harmony" and "fairness" (which are foundational to ethical good) can be conceptualized as pathways to "beauty" and from "love." Toward a pursuit of harmony, a system acquires a sense of beauty. And from a motivation and commitment of love, acts of fairness, that do not necessarily achieve equality, are ethically optimal.

Beauty has been much debated through the millennia. The ancient Greeks consider it one of the three hallmarks of philosophy along with truth and goodness. "Truth" has been embedded within the e-DNA model as accuracy that can be trusted to present an identity that is mutually agreed upon ("desired"). "Goodness" is seen as the DNA of ethical reasoning that included both the good-evil reasoning continuum as well as the full nine overlapping continuums interactively engaged. But "beauty" must be unpacked more intuitively.

The culturally impacted construct of beauty yields a broad diversity that must account for tastes in spatial presentations (e.g. clothing fashions, facial shapes), character generalizations (e.g. virtuous character and beautiful personality) and even beauty in power disruptions (e.g. distant stars forming and exploding). Without a sense of beauty and its opposite, ugliness, ethical reasoning might degenerate to only quantitative measurements of "fairness" and "balance." However, humanity's attraction to beauty and aversion to the violation of beauty (resulting in ugliness) makes ethical reasoning far more interesting and problematic.

Beauty can be perceived through the lens of creative harmony of jealousy space. Beauty can be conceived as displaying creative harmony amidst space that is jealousy held. Similarly, ugliness can be posited as displaying disharmonious jealous space. The good and evil of beauty and ugliness is thus a matter of negotiating jealous space.

The link between beauty and jealous space is intuitive. Beauty reveals an emotional attraction focused on some spatial object or spatially grounded concepts (such as symmetry). An attraction can be conceptualized as a jealousy—a desire to possess for oneself (at some distance) that which is deemed precious. Space that is jealously possessed and is in creative harmony with other jealously held spaces may be deemed beautiful within a family, a business system, a culture. However, when a space is jealousy possessed by conflicting parties, these jealousies (i.e. destructive envies) can produce an ugliness that can lead to brutal conflicts. Thus the underlining dynamics of jealous space is intrinsically embedded within human reasoning of beauty-ugliness.

This beauty is on a continuum with ugliness. Degrees of beauty are compared with degrees of ugliness. Consistent with the above definition of beauty, ugliness is posited as the violation of creatively harmonious jealous space—thus disharmony of envied space. The comparative difference is primarily within the definitions of jealousy and envy. Jealous is a jealousy "for" something or someone with an established right of ownership, while envy is a jealousy "of" something or someone with no established right of ownership. (Obviously, establishing rightful ownership can be problematic.)

For instance, societies agree that parents have some limited right of ownership to their children. For a parent to be jealous "for" the space of his/her child is a beautiful act of harmony. However, when a parent becomes jealous "of" (envious of) the child, something very different occurs, something very ugly. To be jealous "of" is an intrusion of personal space. Parental jealousy "for" can nurture the child while envy, jealousy "of", can rob the child of the space necessary for protection and development.

The desire (and sometimes an act) to invade the space of another and take from him/her that which he rightfully possesses is an ugliness that humanity is acquainted with. This envy, this over-possessive, misdirected and deformed jealousy, can undermine human relationships while a proper sense of jealousy "for" another can help protect and develop a person who is cherished within that possessive jealousy.

For example, if one is jealously possessives of his/her own sexual space (body) and someone attempts to enter that body space without permission, then an internal emotional reaction will occur indicating that this intrusion is an unfair violation, that this act is an ugliness warranting the labels of "evil". Thus, it is culture-general to discuss and condemn the destructive ugliness of sexual rape. Rapes in wartime have sometimes been justified throughout history as acts of conquering the enemy. Fortunately, such wartime violations of jealous space are condemned by the Geneva Convention.

Another common word associated with ethical reasoning is "love". Love is determined by individuals and societies to be both a high ideal and a base passion. Love as an ethic is nebulous. Love may motivate many ethical pursuits. Moreover, the absence of love, when love is expected or longed for, or the presence of hate (love's opposite), invokes ethical choices. Love can be conceptualized as an internal working of beauty and for beauty. And beauty, creative harmony of jealous space, is an outward evidence of some love in action. Furthermore, love as a motive helps mitigate the necessity of fairness that is not always equal.

A final example of the utility of "creative harmony of jealous space" that defines beauty is a tragedy of ugliness and evil. Cruel ugliness reigned in the Rwandan genocide of 1994 in which an estimated 800,000 people were killed in 100 days. One people group, the majority Hutus, sought creative harmony for their desired identity by denigrating their opposition as "cockroaches" (an undesired identity) and systematically labored to eradicate them. They negotiated their space (i.e. their country with physical land and property) with a jealousy that became envious, over-possessive and oppressive to the minority Tutsi population. This negotiation of jealous space allowed a justification for the evils of genocide—a justification acceptable at that time to many (not all) Hutus while being totally unacceptable to all Tutsi. Thus, the e-DNA model can be used in parsing highly charged and ethically implicit behaviors that are disastrously ugly and evil.

The construct of beauty as creative harmony of jealous space holds promise as an e-DNA in negotiating the abstract and practical ethical discussions of our day across cultural distinctions. In going forward, an analysis of ethical reasoning patterns across cultures is needed. This analysis can serve to reinforce the case for this e-DNA model driven by beauty as creative harmony of jealous space.

5 Generality of e-DNA Suggested

The e-DNA model is a generalization that can be useful across various a wide variety of cultural setting. From this generalized model, differences from culture, gender, age, etc. that are prevalent in ethical reasoning can be derived. Generality is suggested through five perspectives.

First, the concept of "creative" can be viewed as a generalization since "change over time" (necessary for creativity) is inherent in all ethical systems of thought—even as language itself changes over time. Second, "harmony" can be perceived as a general ethical construct since its complete opposite insures annihilation of any set of identities (e.g. the destruction of civilizations). Third, jealousy can be projected across cultures from the play of jealousy within the Oedipus complex that has been documented in over 100 cultures (Johnson, A. W. & Price-Williams, D., 1996). Fourth, spatial constructs are inherent in all language at various level of abstraction. Language development starts with objects (e.g. "mommy"), usually associate with some time marker and then over time generalizations and abstractions are formed that make transmission of meanings between persons a fruitful enterprise.

The fifth perspective for suggesting generality will be a specific parsing of a Japanese word, *amae*, using the e-DNA model that has been put forth in English (see Table 1). This cross-cultural evaluation will contribute evidence for the generality of the model.

e-DNA Model	Japanese *Amae* Parsed
Logic of Intellect	
Powerful – powerless	*Amae* requires the powerlessness of receiving as a child would and yields the power of being provided for.
Good – evil	*Amae* requires an acknowledgement of good in one's in-group and holds that evil is betrayal of one's in-group.
Accuracy – intuition	*Amae* requires intuition to negotiate relationships and assumes the accurate interpretation of *amae* as a social construct.
Space	*Amae* requires the negotiation of space between two or more people.
Logic of Emotion	
Trust – fear	*Amae* requires trust in other(s) and it implies the fear of being betrayed by others.
Honor – shame	*Amae* requires the honor of submitting to another's will and it forbids the shame of betraying another.
Freedom – bonding	*Amae* requires the bonding of dependency and yields the freedom of dependency.
Jealousy	*Amae* requires the management of a privileged and thereby jealous relationship between people.
Imagined Outcomes	
Surviving – thriving	*Amae* views the proper networking of relationships for both surviving and thriving.
Desired identity – undesired identity	*Amae* views self as dependent as a desired identity and views the absence of a dependent relationship as an undesired identity.
Meaningful – meaningless	*Amae* views the parent-child relationship as the fundamental meaningful relationship and the absence of *amae* as fundamentally a meaningless existence.
Creative harmony	*Amae* requires both persons in an *amae* relationship maintain and creatively enhance harmony

Table 1: Using e-DNA Model to Parse the Japanese *Amae* Construct

Japanese psychiatrist Takeo Doi (1981) described in detail the dynamics of *amae* in the Japanese culture stating, "The Japanese term *amae* refers, initially, to the feelings that all normal infants at the breast harbor toward the mother – dependence, the desire to be passively loved, the unwillingness to be separated from the warm mother-child circle and cast into a world of objective 'reality' " (p. 7). He went on to say, "… all the many Japanese words dealing with human relations reflect some aspect

of the *amae* mentality. This does not mean, of course, that the average man is clearly aware of *amae* as the central emotion in ninjo (human feeling)" (p. 33).

Regarding the impact of *amae* on the culture, he stated, "Only a mentality rooted in *amae* could produce a people at once so unrealistic yet so clear-sighted as to the basic human condition; so compassionate and so self-centered; so spiritual and so materialistic; so forbearing and so willful; so docile and so violent" (p. 9).

Furthermore, Doi compared the Japanese with Westerners in stating, "Scholars have put forward many different theories concerning the ways of thinking of the Japanese, but most agree in the long run that, compared with thought in the West, it is not logical but intuitive" (p.76). Doi proposed outsiders struggle with the *amae* construct. He stated, "… to persons on the outside who do not appreciate *amae* the conformity imposed by the world of *amae* is intolerable, so that it seems exclusivist and private, or even egocentric" (p. 77).

The e-DNA model analysis of the Japanese construct of *amae* is not intended to fulfill the richness of the Japanese construct but rather to approximate its construction in such a way that the multi-variable applications of *amae* may be anticipated and appreciated within the Japanese cultural context. This analysis of *amae* contributes evidence for the generality of the e-DNA model across human cultures.

6 Using e-DNA in AGI

In hierarchical structures, one would need to prioritize the three proposed central constructs of e-DNA. However as previously mentioned, Wittgenstein suggested "Language is a labyrinth of paths" (Philosophical Investigations 203). This e-DNA model, with overlapping and interacting continuums, accounts for the inherent convolutions—labyrinth of paths—of common human language without establishing a true hierarchy among the central constructs.

Earlier the question arose of accounting for differences in language translations of the words used on the continuums. The labyrinth of paths in language helps alleviate this problem. The assumption that language is discrete and static requires fixed constants that provide exact translations rather than variables within an approximated range. (This range does not allow non-sensical relativism that would cancel the prospect of transference of meanings). This e-DNA model opts for an approximated range of meanings.

The geometrical structures of the e-DNA model lend themselves to computer programming. This set of (3) 3-D grids provides an acceptable means for mapping ethically constructs.

By parsing (with the inputter's bias accounted for) an abundance of words in sentence and image contexts, a more general understanding of the ethical use of a word can be extrapolated. This extrapolation can then be used in evaluating and/or forming ethical rules of thumb. That ethical evaluation would be on a continuum from optimal, acceptable, warning to dangerous.

This e-DNA model can evaluate and suggest optimizing pathways for the richness of ethical reasoning required for true AGI. And without which the imagined outcome of super-human artificial intelligence can only be seen as devastating for humanity. If AGI machines advance with only an ethic of effectiveness and efficiency (inherent in almost all programming), then thriving and surviving might well dominate the struggle between humanity and machine in the decades ahead—with machine the predictable winner.

7 Conclusion

This paper put forth a means of describing an ethical DNA and illustrated its utility in parsing an ethically implicit Japanese construct. In seeking to establish an e-DNA model, I have posited nine

overlapping and interacting continuums with three central constructs. Evidence for its generality has been provided.

If human-level artificial intelligence is to be achieved, the DNA code of thought and behavioral decisions must also be articulated and translated into machine language process and output decisions. Decisions are foundational to human intelligence. The human mind seems to parse all decisions in a seamless fashion while seeking congruence and abating dissonance. This parsing process is mostly opaque to us all.

Describing process (thought) decisions and output (behavioral) decisions are essential for achieving human-level artificial intelligence.

This paper stops short of the point where usage of e-DNA can account for all process and output decisions that human-level intelligence achieves. Future work can extend the e-DNA model to account for this full range of decisions needed for AGI.

References

Asimov, I.: *I, Robot,* New York: Doubleday & Company. (1950)

Clanton, G.: A sociology of jealousy. In G. Clanton & L. G. Smith (Eds.). *Jealousy* (3rd ed.). New York, NY: University Press of America. (1998) 297-312

Doi, T.: *The anatomy of dependency: The key analysis of Japanese behavior*. Tokyo, Japan: Kodansha International (1981).

Ennis, R.: A theoretical model for research in intercultural decision making. *Intercultural Communication Studies.* 8: (2004) 113-124

Gubrud, M.: Nanotechnology and International Security, *Fifth Foresight Conference on Molecular Nanotechnology* (November 1997)

Johnson, A. W. & Price-Williams, D.: *Oedipus ubiquitous: The family complex in world folk literature.* Stanford, CA: Stanford University Press (1996)

Lewis, H. Shame, repression, field dependence, and psychopathology. In *Repression and dissociation: Implications for personality theory, psychopathology and health*. Chicago, IL: University of Chicago Press. (1995) 239-241

Mikhail, J.: Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences.* 11(4) 143-152 (2007)

Pana, L.: Artificial intelligence and moral intelligence. *TripleC* 4(2) 254-264 (2006)

Smith, C., Blake. P., Harris, P.: I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *Plos One* 10.1371 (2013)

Wittgenstein, L.: *Philosophical Investigations*. Trans. G. E. M. Anscombe. Oxford: Blackwell Publishers. (1999)

Reference for paper:

Ennis, R. Ethical DNA Model for Artificial General Intelligence.
USB Proceedings The 10th International Conference on Modeling Decisions for Artificial Intelligence. MDAI 2013, Barcelona, Catalonia November 20 - 22, 2013, Pages 56 -67.
ISBN: 978-84-695-9120-8 D.L.: B. 27203-2013

Evaluating Virtue and Vice Values for Ethical Artificial General Intelligence

Ralph C. Ennis
Raleigh, NC USA
ralphennis@gmail.com

Abstract. This paper employs an ethical DNA model (Ennis, 2013) to map virtue and vice values commonly accepted within the Western world and a virtuous Japanese construct. Plato posited *temperance* as one of four cardinal virtues by which all other virtues could be established. The vice of *envy* is one of seven capital vices highlighted by the Catholic Church. Both temperance and envy are parsed using the eDNA model. Similarly, the virtuous Japanese construct *amae* will also be parsed to demonstrate the utility of the model across cultural differences. Subsequently, a mean location with intensity of attraction and aversion regarding *amae* will be mapped onto a two-dimensional graph. Then, using an evaluative grid, an assessment of that placement will suggest the ethical acceptability of the *amae* construct. After accounting for bias, this evaluating process of value-laden constructs is adaptable for coding and for establishing a way forward for ethical evaluation and learning within artificial general intelligence.

Keywords. Ethics, artificial general intelligence, virtue, vice

1 Introduction

Across the artificial general intelligence community, there is a growing awareness of the need for embedded ethical reasoning in AGI. This need is deemed urgent as artificial agents become more pervasive (Shulman, et.al. 2009). And Geortzel and Pitt (2012) suggest a co-evolution of AGI and AGI-related ethical theory as a convergent way forward to meet this need.

Moor (2006) has differentiated between implicit and explicit ethical agents. Implicit agents are programmed to behave ethically. Explicit agents are programmed to use ethical principles. Furthermore, Anderson and Anderson (2007) confirm that explicit agents are the ultimate goal for machine ethics.

Goertzel and Bugaj (2007) have used various models of cognitive and ethical development (such as Piaget, Perry and Kohlberg) to project possible stages of ethical development in AGI systems. In addition, they posit ethical imperatives for human-AGI interaction (Bugaj and Goertzel (2007). Together these stages and imperatives form a means of projecting the development of ethical AGI.

Furthermore, Potapov and Rodionov (2012) suggest that hierarchical value learning rather than reward maximization is crucial for AGI to be safe. Rewards, they state, are not to be valued for themselves but rather the values they reward are to be reinforced.

In this paper, I will use an ethical DNA (eDNA) model for AGI (Ennis, 2013) to parse and map values that are generally accepted as virtues and vices and that interlace ethical principles. Plato's cardinal virtue of *temperance*, the capital vice *envy* of the Catholic Church and the Japanese construct of *amae* will be used as values. No hierarchy of these values will be offered. Rather the parsing of each of these values can be mapped onto a grid for evaluation of ethical acceptability. The mapping and evaluation of *amae* will serve to illustrate this utility.

2 Plato's Virtue of Temperance

Plato posited four cardinal virtues as key to human society. These are temperance, prudence, courage (fortitude) and justice. These virtues are cardinal in that from these all other virtues are perceived by Plato to have their basis. For this paper only one virtue will be parsed—temperance. Temperance is commonly defined (Wikipedia) as "moderation in action, thought of feeling; restraint."

Table 1 displays a parsing of temperance using an ethical DNA model (Ennis, 2013). This parsing suggests that other virtues can be deconstructed by the nine continuums and three central constructs of the eDNA model.

eDNA Continuums	Temperance
Logic of Intellect	
Power - Powerless	Temperance requires the power of self-control in the face of temptations to indulge.
Good - Evil	Temperance is perceived as a good quality and practice.
Accuracy - Intuitive	Temperance is a fuzzy concept. The limits for being non-temperate is often difficult to precisely define.
Space (as a central construct of intellect)	Temperance implies spatial constructs of what one is temperate for.
Logic of Emotion	
Trust - Fear	Temperance requires trust in the face of fear of loss.
Honor - Shame	Temperance often brings a sense of honor that one is not controlled by one's desires. Intemperance also brings shame.
Freedom - Bonding	Temperance brings freedom from one's desires.
Jealousy (as a central construct of emotion)	Temperance implies a jealousy for that which is a better long-term gain vs. a jealousy of (envy) that which is at hand.
Imagined Outcomes	
Thriving - Surviving	Temperance can improve one's chances of thriving.
Desired Identity - Undesired	Temperance can be a desired identity as in "I am a temperate person."
Meaningful – Meaningless	Temperance implies that life has a meaning apart from immediate fulfillment of desires.
Creative Harmony (as a central construct of imagined outcomes)	Temperance seeks to create a harmony within one's self.

Table 1: Using an eDNA Model to Parse Plato's Cardinal Virtue of Temperance

3 Catholic Church Vice of Envy

In similar fashion, vices—the antithesis of virtues—can be mapped using the same eDNA framework. Table 2 is a parsing of the vice *envy* which is eschewed by the Catholic Church.

Envy is commonly defined (Wikipedia) as a resentment that "occurs when someone lacks another's quality, achievement or possession and wishes that the other lacked it."

eDNA Continuums	Envy
Logic of Intellect	
Power - Powerless	Envy assumes a powerless state in pursuit of power.
Good - Evil	Envy is mostly perceived as an evil.
Accuracy - Intuitive	The boundaries of envy are mostly intuitive.
Space	Envy is played out in other's space.
Logic of Emotion	
Trust - Fear	Envy is a fear of unmet longings.
Honor - Shame	Envy is mostly shameful.
Freedom - Bonding	Envy is a bondage seeking a freedom.
Jealousy	Envy is a jealousy of someone's better position or possessions.
Imagined Outcomes	

Thriving - Surviving	Envy seeks to thrive at another's expense.
Desired Identity – Undesired	Envy is an undesirable identity except through shamelessness.
Meaningful – Meaningless	Envy is mostly meaningless.
Creative Harmony	Envy seldom creates harmony.

Table 2: Using the eDNA Model to Parse Catholic Vices

4 Japanese *Amae*

The eDNA model can be useful across cultural setting. This utility is suggested by the parsing of a Japanese word, *amae*, using the eDNA model (see Table 3). As previously stated (Ennis, 2004, 2013):

> Japanese psychiatrist Takeo Doi (1981) described in detail the dynamics of *amae* in the Japanese culture stating, "The Japanese term *amae* refers, initially, to the feelings that all normal infants at the breast harbor toward the mother – dependence, the desire to be passively loved, the unwillingness to be separated from the warm mother-child circle and cast into a world of objective 'reality' " (p. 7). He went on to say, "… all the many Japanese words dealing with human relations reflect some aspect of the *amae* mentality. This does not mean, of course, that the average man is clearly aware of *amae* as the central emotion in ninjo (human feeling)" (p. 33).
>
> Regarding the impact of *amae* on the culture, he stated, "Only a mentality rooted in *amae* could produce a people at once so unrealistic yet so clear-sighted as to the basic human condition; so compassionate and so self-centered; so spiritual and so materialistic; so forbearing and so willful; so docile and so violent" (p. 9).
>
> Furthermore, Doi compared the Japanese with Westerners in stating, "Scholars have put forward many different theories concerning the ways of thinking of the Japanese, but most agree in the long run that, compared with thought in the West, it is not logical but intuitive" (p.76). Doi proposed outsiders struggle with the *amae* construct. He stated, "… to persons on the outside who do not appreciate *amae* the conformity imposed by the world of *amae* is intolerable, so that it seems exclusivist and private, or even egocentric" (p. 77).
>
> The eDNA model analysis of the Japanese construct of *amae* is not intended to fulfill the richness of the Japanese construct but rather to approximate its construction in such a way that the multi-variable applications of *amae* may be anticipated and appreciated within the Japanese cultural context. This analysis of *amae* contributes evidence for the generality of the eDNA model across human cultures.

eDNA Continuums	**Japanese *Amae* Parsed**
Logic of Intellect	
Powerful – powerless	*Amae* requires the powerlessness of receiving as a child would and yields the power of being provided for.
Good – evil	*Amae* requires an acknowledgement of good in one's in-group and holds that evil is betrayal of one's in-group.
Accuracy – intuition	*Amae* requires intuition to negotiate relationships and assumes the accurate interpretation of *amae* as a social construct.
Space	*Amae* requires the negotiation of space between two or more people.
Logic of Emotion	
Trust – fear	*Amae* requires trust in other(s) and it implies the fear of being betrayed by others.

Honor – shame	*Amae* requires the honor of submitting to another's will and it forbids the shame of betraying another.
Freedom – bonding	*Amae* requires the bonding of dependency and yields the freedom of dependency.
Jealousy	*Amae* requires the management of a privileged and thereby jealous relationship between people.
Imagined Outcomes	
Surviving – thriving	*Amae* views the proper networking of relationships for both surviving and thriving.
Desired identity – undesired identity	*Amae* views self as dependent as a desired identity and views the absence of a dependent relationship as an undesired identity.
Meaningful – meaningless	*Amae* views the parent-child relationship as the fundamental meaningful relationship and the absence of *amae* as fundamentally a meaningless existence.
Creative harmony	*Amae* requires both persons in an *amae* relationship maintain and creatively enhance harmony

Table 3: Using the eDNA Model to Parse the Japanese Construct of *Amae*

Mapping *amae* can be achieved by assigning a number on a sliding scale across each of nine continuums with endpoints (See Figure 1). In addition, an intensity scale is added to account for degrees of attraction and aversion to each virtue or vice.

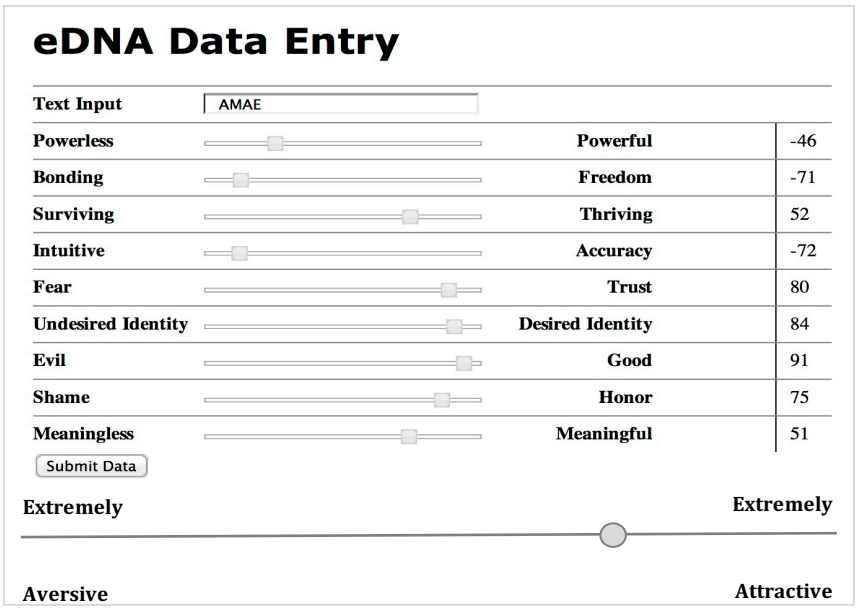


Figure 1 Sample input for *amae*

In Figure 2 the assigned scales are mapped at the assigned intensity of attraction-version. A mean (in black) is derived for *amae* at the assigned intensity.

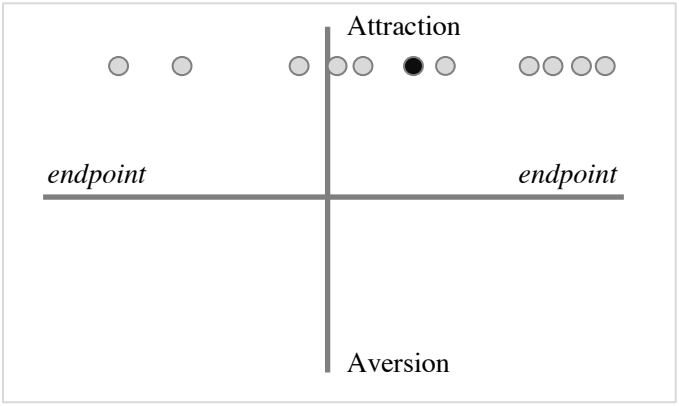


Figure 2 Map of *amae* sample input with mean derived in black

5 Evaluative Grid

In order to assess the ethical acceptability of a mapped construct, an evaluative grid is posited. This grid will assess ethical acceptability on a scale from optimal, acceptable, and warning to dangerous. In order to establish this grid, the endpoints are classified for desirability.

Each endpoint of the nine continuums can be classified as either desirable or significantly less desirable. Desirable endpoints are generally pursued by both individuals and across cultures. Those endpoints are power, accuracy, good, honor, trust, freedom, thriving, meaningful, and desired identity. These endpoints will be noted as "Class A" endpoints.

Class B endpoints are generally less desirable than Class A endpoints. Class B endpoints include powerlessness, intuitive, evil, shame, fear, bonding, survival, meaningless, and undesired identity.

Figure 3 details an evaluative grid for ethical acceptability. And Figure 4 plots the input of *amae* from Figure 2 onto that grid. From this sample input *amae* is perceived by an individual inputter to be on the edge of "acceptable" and "warning".

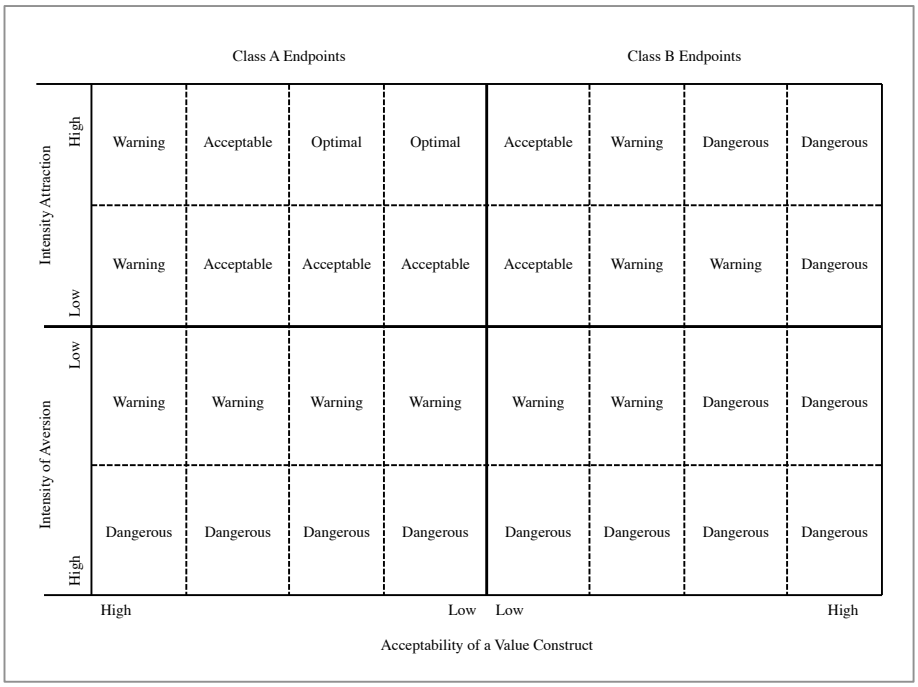


Figure 3 Evaluative Grid for Ethical Acceptability

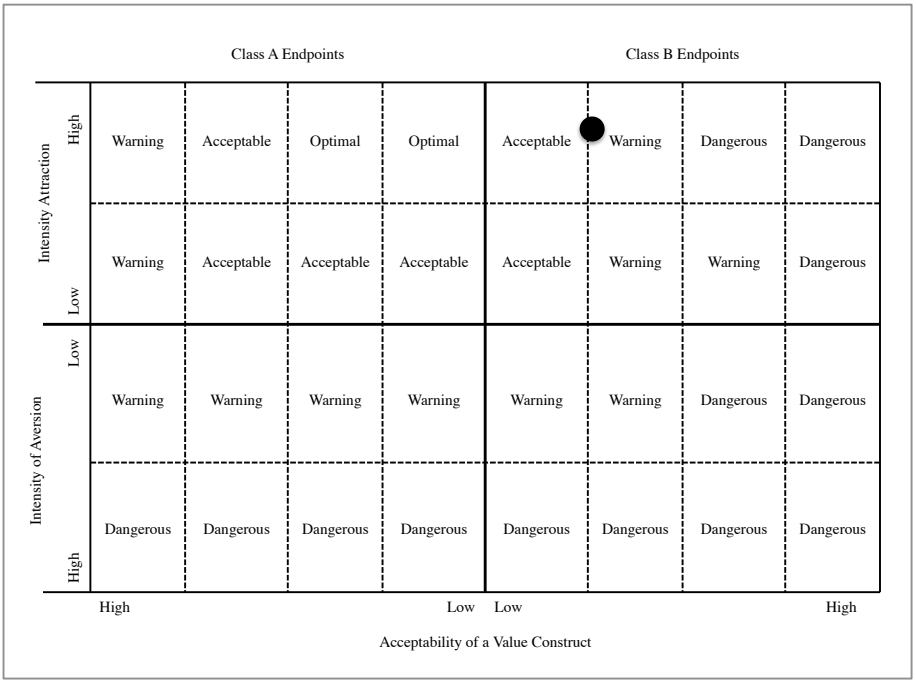


Figure 4 Assessment of Ethical Acceptability from *Amae* Sample Input

The above mapping and evaluation process can be used for individual words or value-laden statements in order to assess ethical acceptability. Future work includes enhancing the accuracy of the grid through statistically validating the initial acceptability assignment of each section of the grid.

Since individuals and cultures vary in their perception of ethical acceptability of each of the above virtue and vice values (and all value constructs), a means for accounting for ethical bias is essential. A bias factor can be established for inputters by mapping their preferences regarding the eDNA continuums. A future paper will posit a mechanism to account for bias.

As values are inputted and evaluated, patterns of ethical acceptability may surface. These patterns can shape future evaluations and help establish rules of thumb for ethical reasoning within artificial general intelligence agents.

6 Conclusion

By parsing a cardinal virtue of Plato, a Catholic Church vice and the Japanese construct *amae*, I am suggesting the eDNA model (Ennis, 2013) can not only serve as a means of mapping complex value constructs that are pervasive within human society but, with the addition of the evaluative grid described in this paper, can also provide a way of assessing the ethical acceptability of constructs. And with additional mapping inputs and assessments, patterns may form to better test the ethical acceptability of any value-laden proposition.

This mapping process and evaluative grid can be coded. Thus, with a programmable means of assessing ethical acceptability of complex constructs, a way forward for developing ethical learning and reasoning for AGI agents can be explored.

References

Anderson, M. and Anderson, S.L.: Machine ethic: Creating an ethical intelligent agent. *AI Magazine* 28:4 (2007) 15-25

Bugaj, S. V. and Goertzel, B.: Five ethical imperatives and their implications for human-AGI interaction. Novamente LLL and AGIRIR Institute. (2007)

Cardinal Virtues. http://en.wikipedia.org/wiki/Cardinal_virtues (2014)

Doi, T.: *The anatomy of dependency: The key analysis of Japanese behavior*. Tokyo, Japan: Kodansha International (1981).

Ennis, R.: A theoretical model for research in intercultural decision making. *Intercultural Communication Studies.* 8: (2004) 113-124

Ennis, R. Ethical DNA model for artificial general intelligence.
USB Proceedings The 10th International Conference on Modeling Decisions for Artificial Intelligence. MDAI 2013, Barcelona, Catalonia November 20 - 22, 2013, Pages 56 -67. ISBN: 978-84-695-9120-8 D.L.: B. 27203-2013

Goertzel, B. and Bugaj, S. V.: Stages of ethical development in artificial general intelligence systems. Novamente LLL and AGIRIR Institute. (2007)

Goertzel, B. and Pitt, J. Nine ways to bias open-source AGI toward friendliness.
Journal of Evolution and Technology. Vol. 22 Issue 1 (February 2012) 116-131

Moor, J. H. 2006. The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems* 21(4): 18–21.

Potapov, A. and Rodionov, S.: Universal empathy and ethical bias for artificial general intelligence. http://arxiv.org/abs/1308.0702 (2012)

Seven deadly sins. http://en.wikipedia.org/wiki/Seven_deadly_sins (2014)

Shulman, Carl, Henrico Jonsson, and Nick Tarleton. 2009. "Machine ethics and superintelligence." In *AP-CAP 2009: The Fifth Asia-Pacific Computing and Philosophy Conference, October 1st-2nd, University of Tokyo, Japan, Proceedings*, edited by Carson Reynolds and Alvaro Cassinelli, 95–97.

Paper submitted to AGI 2014 Conference on February 18, 2014

Thought Dynamo Decision Mapping Model

2009
Updated May 2013

Abstract

In establishing a methodology for mapping neuron activity and memory within mice onto a 3-dimensional graphic, Tsien (2007) has suggested, by inference, the possibility of mapping human thought onto a 3-D interpretative grid. This paper will set forth a 3-D interpretative model for mapping 'rules of thumbs' for thought and decision within humans. That model will be referred to as the Thought Dynamo Decision Model (TDD Model).

The TDD Model includes three overlapping 3-D grids that represent thought from the perspectives of intellect, emotions and imagined outcomes. The continuums of the logic of intellect are power-powerless, good-evil and accuracy-intuitive. The logic of emotion has three continuums: trust-fear, honor-shame, freedom-bonding. And the imagined outcomes continuums are thriving-surviving, desired identity-undesired identity, and meaningful-meaningless. The intersection of each grid forms a central tendency. For the logic of intellect the central tendency is space; for the logic of emotion the central tendency is jealousy; and the central tendency of imagined outcomes is creative harmony. Thus the overlap of central tendencies becomes a creative harmony of jealous space.
A five step process of decision is posited: plot, associate, adjust, solidify and employ. Connectivity between thoughts and decision is apparent. Thoughts are required to make decisions and decisions are made within thinking. The dynamics of thought and decision are geometrically modeled by assigning weights to inputs, strength to associations, dynamically accounting for adjustments and solidifications over time, and projecting rules of thumb in decision making onto (3) 3-D axes.
This model of thought and decision involves emotionality across cultural differences. The significant obstacle of diverse emotionality in positing a culture-general model of thought and decision is addressed by examining the Japanese construct of "amae".

The TDD Model can be used to shape research agenda. The research outcomes can be descriptive, predictive and prescriptive in nature.

Key Words: Decision, intellect, emotion, imagined outcomes, thought, mapping

Contents

Introduction 57

Overview of TDD Model 59

Thought Dynamo 61
- Logic of Intellect 62
- Logic of Emotions 62
- Imagined Outcomes 62
- Overlapping Dynamics 62
- Expanded Continuum 63
- An Example of Movements 65

Decision Dynamics 67
- Plot 67
- Associate 68
- Adjust 68
- Solidify 69
- Goal Seek 70
- Employ 72

Mapping 75
- Rules for Optimization 79
- Mapping the Japanese Construct 'Amae' 80
- Computer Modeling 82
- Future Research 82

Conclusion 83

References 83

Introduction

Decisions are influenced by rational, emotional and imaginative thoughts. Various decision models have influenced the construction of the thought and decision model within this paper. That model will be referred to as the Thought Dynamo Decision Model (TDD model). Aspects of the Expected Utility Model, Prospect Model, Self-Regulation Model, and Reciprocal Causal Model are found within the TDD model.

The Expected Utility Theory (Fishburn, 1970) focuses on what action should be taken in a particular situation – what is the best choice. That determination of "should" or "best" is based firmly upon rationally exploring relevant data and options with some anticipation of an expected (imagined) outcome. Prospect theory (Tversky and Kahneman, 1974) addresses heuristics principles, rules of thumb which people employ to shorten the decision process. These rules of thumb can cause decision bias as well. The influence of a person's emotionality within prospect theory has been explored by Kuula and Salminen (1996). Byrnes (1988) put forth a self-regulation model of decision making. This model suggests people make decisions as an adaptive process in order to attain their own goals. At every juncture where a goal is blocked, people adapt (self-regulate) their decisions in order to attain their goals. Reciprocal causal models are based on the idea that feedback from decisions impacts the basic harmony of a system and that diversity, not homogeneity, provides more harmony (Maruyama, 1987).

A consideration regarding decision models is the concept of a super-logic premise. This premise suggests that across cultures there is only one right way to logic, all other ways are inferior to super-logic. Maruyama (1987) critiques super-logic and posits that the theory of super-logic is "tautological rather than logical" (p.87).

The TDD model will employ aspects of these theories while formulating a model that has descriptive, predictive and prescriptive value. This model posits a culture-general construct and thus implicitly accounting for logics across cultures.

In order to establish a geometric model of thought and decision, three types of space must be delineated – physical space, neural space and interpretative space (see figure 1).

Physical space is that space perceived by the mind to be external to the human mind but containing all else including neural space within the brain. Physical space can be perceived through sensory input by the mind. This interconnectivity of physical and mental space allows for the philosophical dialogue of a possible real-illusion paradoxical relationship.

Neural space is primarily within the brain. MRI brain research employs mapping techniques to identify areas of the brain that are engaged during specific mental activities. Those mental activities imply an interpretative space.

Interpretative space allows sensory input to be mapped and continuously processed (with varying weights) within a nature-culture schema that can be modeled on (3) 3-D overlapping axes. It is interpretative space that this paper will address.

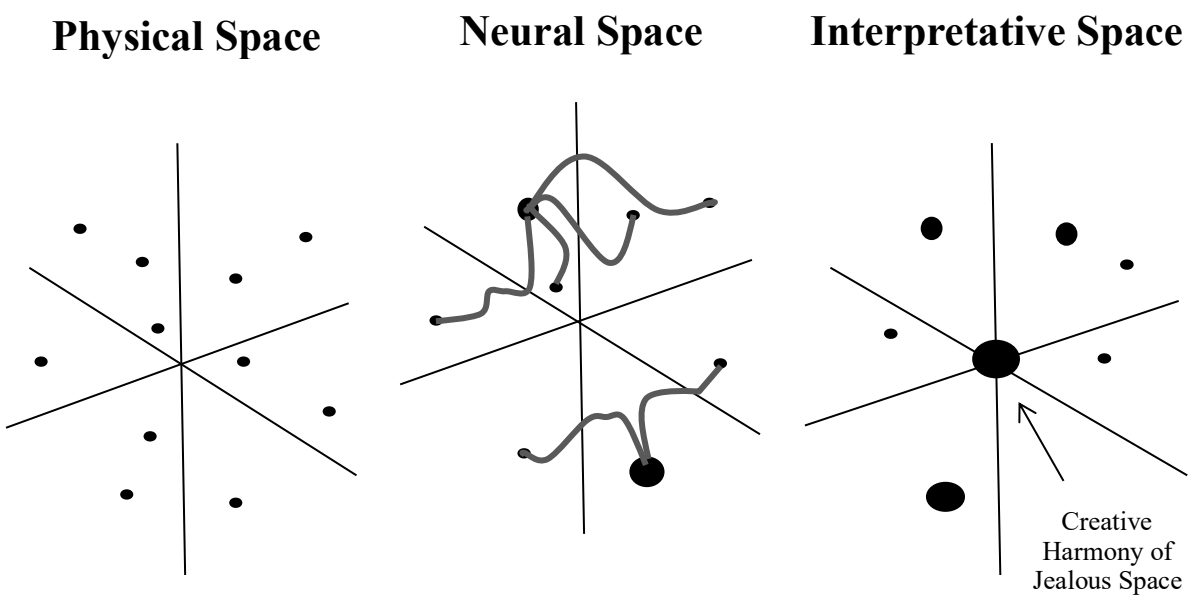


Figure 1: Types of Space

The creator of the smart mouse, Tsien (2007), within the fields of pharmacology and biomedical engineering, has suggested a means for mapping sensory input onto an interpretative space within mice. By inference, similar axes of interpretative space can be conceptualized for humans. The key would be to identify the axes that are most viable for human thought. It is those axes that are posited later in this paper.

> "Software translated the data from an individual mouse into a 3-D plot that represented the activity of the full ensemble of recorded neurons when the animal was at rest and undergoing startling events. Such plots enabled researchers to "read" what was happening to an animal simply by watching the recorded signal move within that 3-D space." (Tsien, 2007, p. 56)
>
> "Remus Osan – another postdoctoral fellow – and I analyzed the recordings using powerful pattern-recognition methods, especially multiple discriminant analysis, or MDA. This mathematical method collapses what would otherwise be a problem with a large number of dimensions (for instance, the activities of 260 neurons before and after an event, which would make 520 dimensions) into a graphical space with only three dimensions." (Tsien, 2007, p. 55)
>
> Space is seen as a panhuman intellectual construct that is useful for examining thought and decisions across cultures. The philosopher Wittgenstein (1958) referred to the "spatial and temporal phenomenon of language" (p. 47). Since all language systems are purported to contain a spatial quality, it can be reasoned that spatial constructs with movements across time is an embedded panhuman thought platform.

Overview of TDD Model

At the center of the TDD model (see figure 2) is an ongoing process, a thought dynamo that can be modeled as three overlapping 3-D axis involving the logic of intellect, logic of emotions and imagined outcomes. At all aspects of this model of thought and decision, this thought dynamo is in action. The thought dynamo will be further examined in the next section.

Decisions, generated by the thought dynamo, can be classified into five broad categories: plot, associate, adjust, solidify, and employ. [These decisions will be examined in more detail within the section "Decision Dynamics".]

1. Plot Decisions: Decisions are made to plot various sensory inputs that come from goal seeking or from non-solicited input or from feedback from outcomes of previous decisions or from thoughts as inputs. Each plotted point is assigned a location and weight. For instance, if a dog walks by a crying baby in a stroller and stops to sniff, our minds will plot many inputs.

2. Association Decisions: The decision to associate (link) input with similar input is also a decision to assign some strength to that association. This decision is often performed with little awareness. For instance, the dog will likely be associated with other inputs of dogs.

3. Adjustment Decisions: New input impacts previous input. Thus with each new plot (location and weight) and association (link and strength) comes a readjustment (of various degrees) of all relevant inputs. If the sniffing dog, from the above example, is a German shepherd, the strength of association may be extremely strong or substantially weak depending on previous experiences with this type of dog. The initially assigned strength will then be adjusted to accommodate the specific context in which the dog is sniffing a helpless baby.

4. Solidification Decisions: As time progresses, decisions are solidified with various degrees of congruence, contradiction, disillusionment and paradox. This solidification forms various types of rules of thumb. If the strength of association of a German shepherd sniffing a baby was weak, as we continued walking past this scene and replayed the scene in our memories, we might solidify our previous association and adjustment decisions as we reinterpreted this memory.

5. Employment Decisions: In the course of living, we make many behavioral decisions. In so doing we employ previously plotted input that is assigned strength of association, adjusted and then solidified using a thought dynamo of logic of intellect, logic of emotions and imagined outcomes. This employment results in verbal and non-verbal actions that impact outwardly.

Each of these decisions are linked to form a very complex system of thought and behavioral decision making across various aspects of time, notably a present, past and future. Plotting sensory input that is "beyond me", i.e. external to my thoughts, and that arrives in a present is influenced by our culturally impacted perceptions and by the awareness of our consciousnesses. At any given moment we are aware of various aspects of input while other aspects "blow through" our mental awareness. That awareness can be culturally honed to "look for" various types of input while ignoring other inputs.

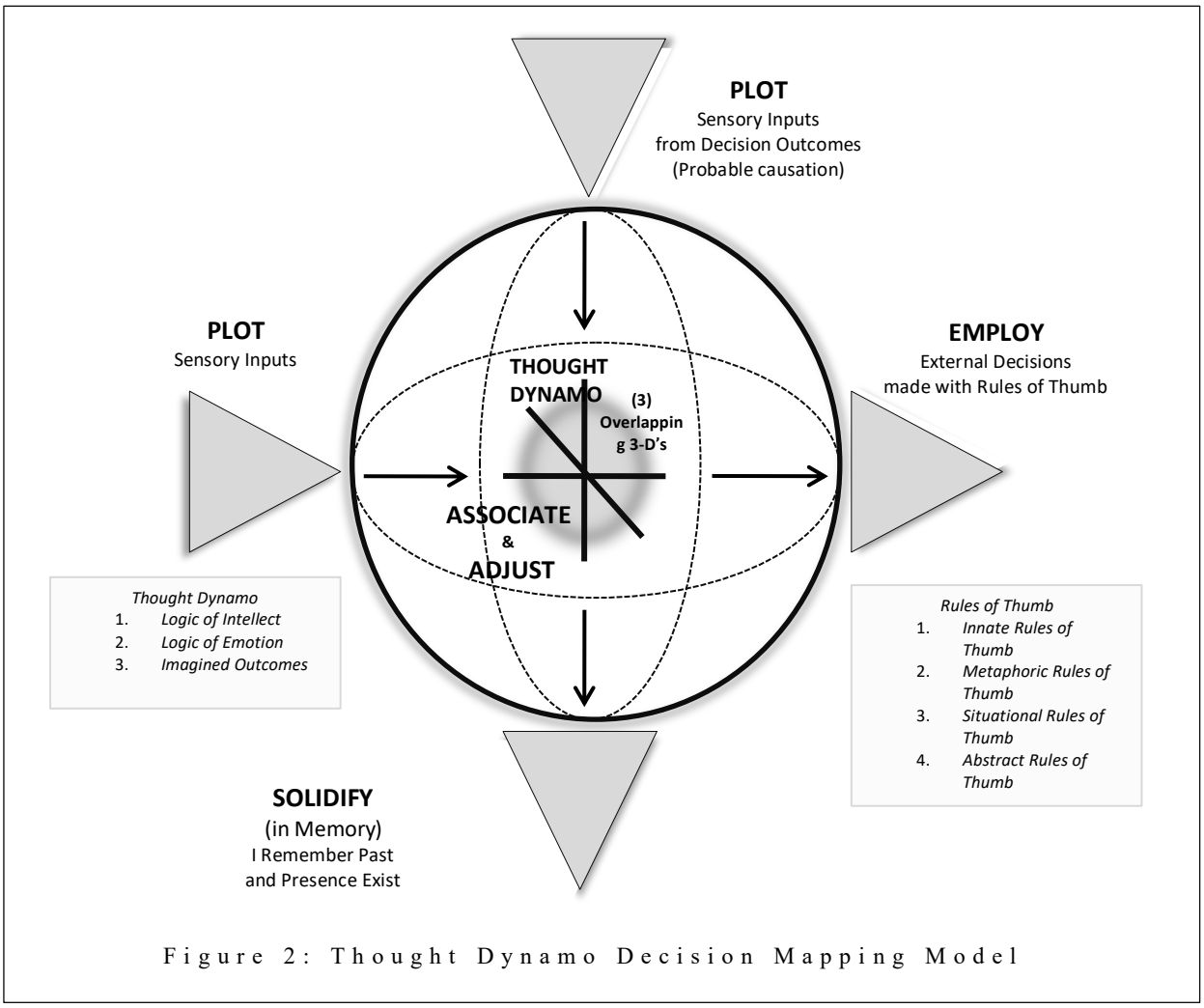


Figure 2: Thought Dynamo Decision Mapping Model

There are three fundamental limitations of the model. First, the model seeks to describe the "software" of thought – not the "hardware" of the brain and yet not contradicting the hardware limitations.

Second, this model does not easily deal with false data and with irony. Straightforward input is more easily handled. However, as more sophisticated rules of thumb are solidified, false and ironical inputs are more accurately addressed.

And third, the model enfolds onto itself from all points into the paradox of "self defining self". The connectivity between thought and decision is apparent and circular. Thoughts are required to make decisions and decisions are made within thinking.

The next two major sections will examine the thought dynamo and decision dynamics. Afterwards the topic of mapping will be addressed along with an example from the Japanese cognitive construct of "amae" (Doi, 1981). This example will support the utility of the model across cultures.

Thought Dynamo

The thought dynamo (see figure 3) is at the center of the TDD model. It drives and is driven by the five types of decisions previously discussed. The model posits three intersecting axes: logic of intellect, logic of emotion and imagined outcomes (Ennis, 2004; Kulich, et al., 2001). These intersecting axes can be conceptualized to account for such abstract notions as beauty, passion and love and amae.

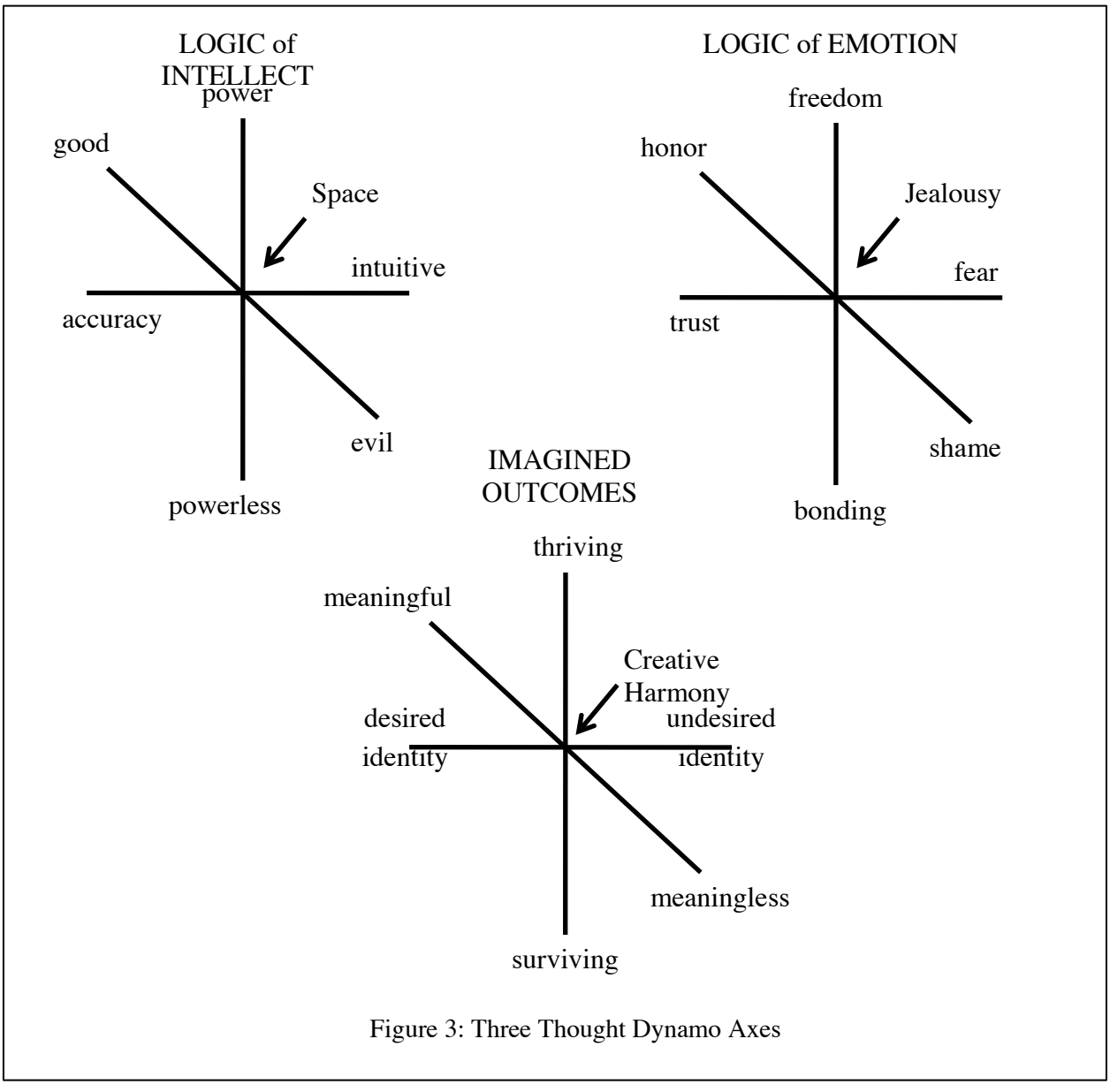


Figure 3: Three Thought Dynamo Axes

Logic of Intellect

The logic of intellect in this model has three axes: power axis, moral axis, and certainty axis (see figure 3). The power axis is a continuum from power to powerless. The moral axis is a continuum from good to evil. The certainty axis is a continuum from accuracy to intuition. And the central construct of the logic of intellect is space. Each of the specifics of the axes is defined by every culture. However, the model posits that all cultural logics involve these three axes.

Logic of Emotions

Similarly, the logic of emotion (see figure 3) is posited to involve three axes: relational, hierarchal and liberty. The relational axis is a continuum from trust to fear. The hierarchal axis is a continuum from honor to shame. The liberty axis is a continuum from freedom to bonding. And the central construct of the logic of emotion is jealousy. Jealousy is a strong emotion "critical to the success and proliferation of our ancestors" (Buss, 2000). And a jealous reaction has been observed in six-month-old infants (Hart and Carrington, 2002). [This assertion is in contradiction to Poulson's (2000) suggestion that shame is the master emotion.]

These few emotional constructs can account for many emotional states. Each of the thirty-four emotions posited in the PAD (Pleasure – Arousal – Dominance) Emotional-State Model (Mehrabian, 1995) can be conceptualized as blended from the three continuums of the logic of emotion.

Imagined Outcomes

The construct of imagined outcomes is posited to involve three axes: adaptation, meanings and identity. The central construct of imagined outcomes is posited as creative harmony (see figure 3).

As a "present" process, the logic of intellect and emotions both have two other time elements embedded in them. The past has been processed and stored through these axes. And a second time element can be conceptualized as imagined outcomes. While acknowledging that the duration of time implicit in imagination is culture-specific, this model positions imagination into a forward spatial non-presence as a culture-general construct. A future imagined outcome emphasizes the possibility of decisions modifying a present assessment of reality. These imagined outcomes also create the possibility of re-interpreting the past with present thoughts and decisions.

Overlapping Dynamics

Figure 4 depicts the intersecting dynamic of the three sets of axes. Good, honor and meaningful interact on a continuum with evil, shame and meaninglessness. Powerful, freedom and thriving interact of a continuum with powerless, bonding and survival. And accuracy, trust and desired identity interact on a continuum with intuition, fear and undesired identity. The central construct of

these intersecting axes is creative harmony of jealous space. The logic of emotion, logic of intellect and imagined outcomes are therefore conceptualized to interact with and influence each other. The relative weights of these factors provide the difference in cultural decision making preferences. It should also be noted that if these axes 'twist', that twisting can account for various aspects of mental health dysfunction.

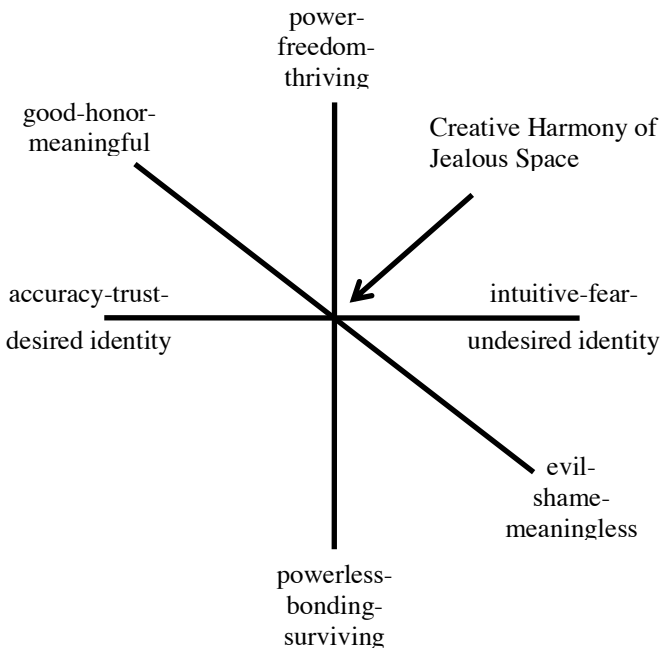


Figure 4: Overlapping Continuums

Expanded Continuum

Each of the nine continuums can be expanded to show the complexity of thought and decision. Figure 8 expands the trust-fear continuum. All other continuum can be similarly expanded. Below is an explanation of this expansion by use of internal "I" thoughts.

Unexamined:
I trust (someone, something) without significantly examining a variety of circumstances.

Reasoned:

I have reasoned with logic of intellect and emotion and imagined outcomes and have concluded that I will trust (someone, something) in general.

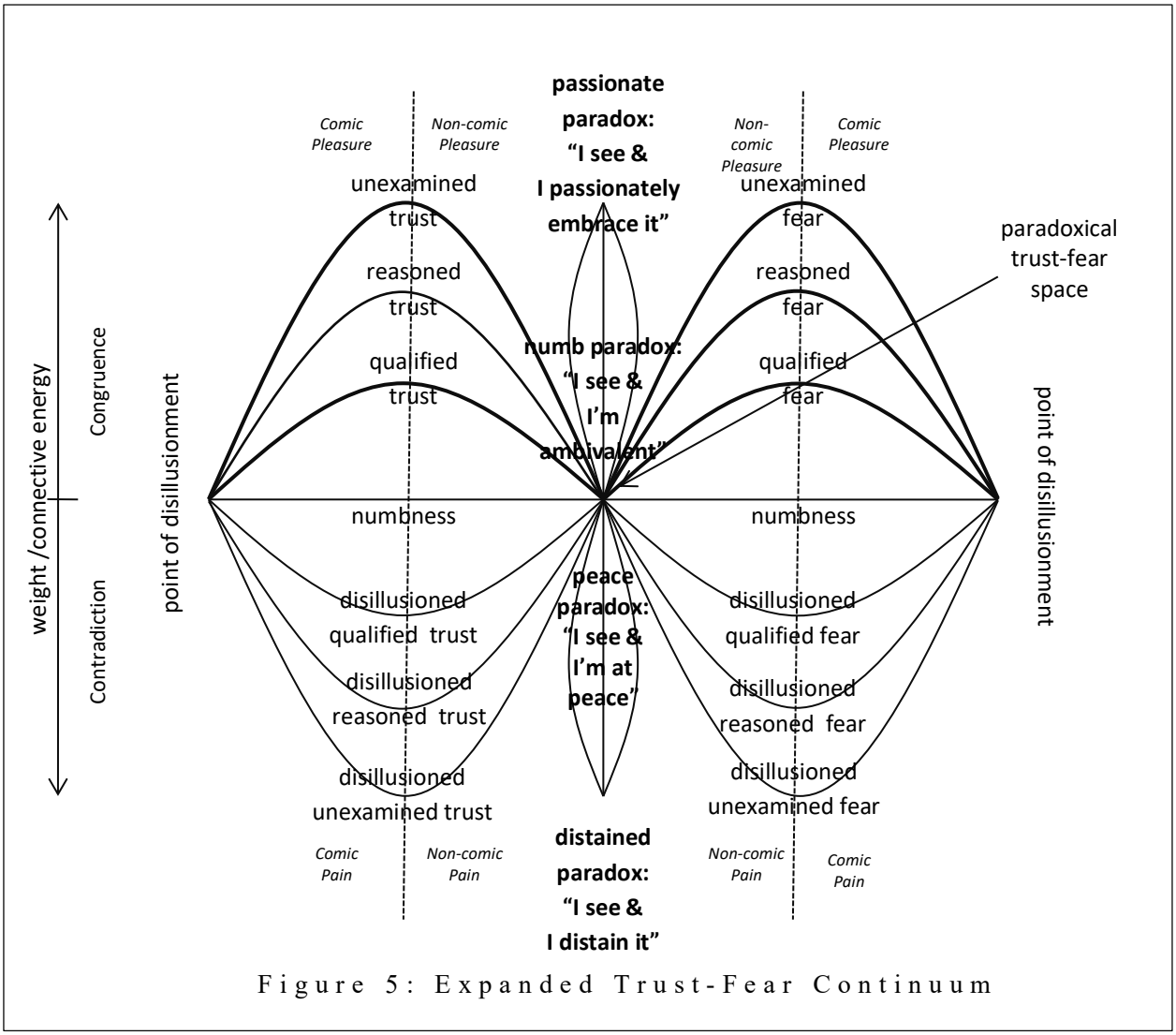


Figure 5: Expanded Trust-Fear Continuum

Qualified:
I will trust (someone, something) under specific circumstances only.

Numbness:
I am ambivalent (numb) regarding this arena of thought.

Disillusionment:
I am disillusioned. Further interpretation of cause and effect has led me to conclude that my original weighted position was inaccurate thus creating cognitive dissonance. I may live with this disillusionment or "flip" it to the other side of the axis (e.g. disillusioned reasoned trust may become reasoned fear) or in extreme cases this may lead to sheering of axes.

Paradox:
I am living with the paradox of trusting what I fear and fearing what I trust. I may passionately embrace this paradox or distain it or be numb (and disconnected) or be at peace (and connected) about it.

An Example of Movements

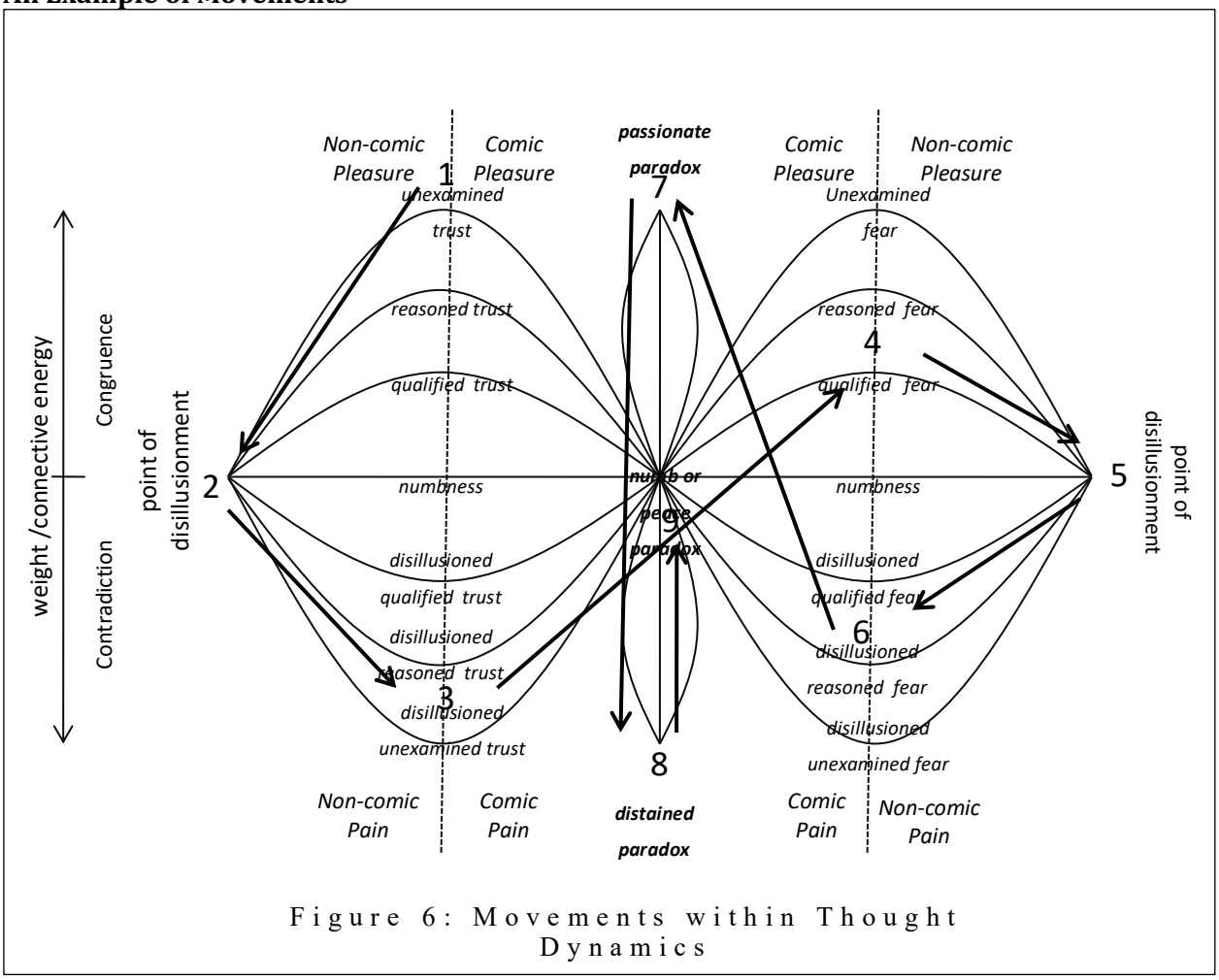


Figure 6: Movements within Thought Dynamics

Below is a description of various possible movements within the thought dynamo (see drawing 8). These movements account for the decision to change one's mind as more input is gathered or more time to process has occurred.

1. One might begin his/her journey of trusting some person from a location of "unexamined trust" (see location 1 on drawing 8).
2. If contrary input (often a non-comic pain type of input such as being lied to) outweighs this "unexamined trust", then a "point of disillusionment" (location 2) might occur.

3. This might move to a point of "disillusioned qualified trust" (location 3) possibly enhanced by a sense of non-comic pleasure (e.g. the pleasure of regaining control in the relationship by moving through disillusionment).
4. "Qualified fear"(location 4) might follow, possibly enhanced by a sense of comic pleasure (e.g. a comic pleasure of feeling that the other was beaten at his/her own game).
5. If sufficient input occurred (often non-comic pain) that outweighed this "qualified fear", then a second "point of disillusionment" (location 5) might occur. (This pain might include a personal sense of shame for not forgiving the person for his/her previous breach of trust.)
6. "Disillusioned reasoned fear" (location 6) might follow and be enhanced by non-comic pleasure of feeling superior to one's previous conclusions.
7. With sufficient input and/or reason, a "passionate paradoxical" (location 7) state might occur. That is, "I see and I passionately embrace" that this particular person can be trusted and feared simultaneously.
8. As time proceeded and the impact of this paradox is absorbed into the decision process, he/she might "distain this paradox"(location 8) as a complexity that doesn't facilitate decision goal seeks.
9. If enough pain (comic or non-comic) occurs, then a sense of unconnected numbness might set in: "I see the paradox and I'm ambivalent" (location 9). This unconnected numbness may be wearisome as the play of comic and non-comic pain and pleasure continues.
10. If enough pleasure occurs, then a sense of connected peace within this paradox might settle in (a sufficiently weighted point at location 9).

Decision Dynamics

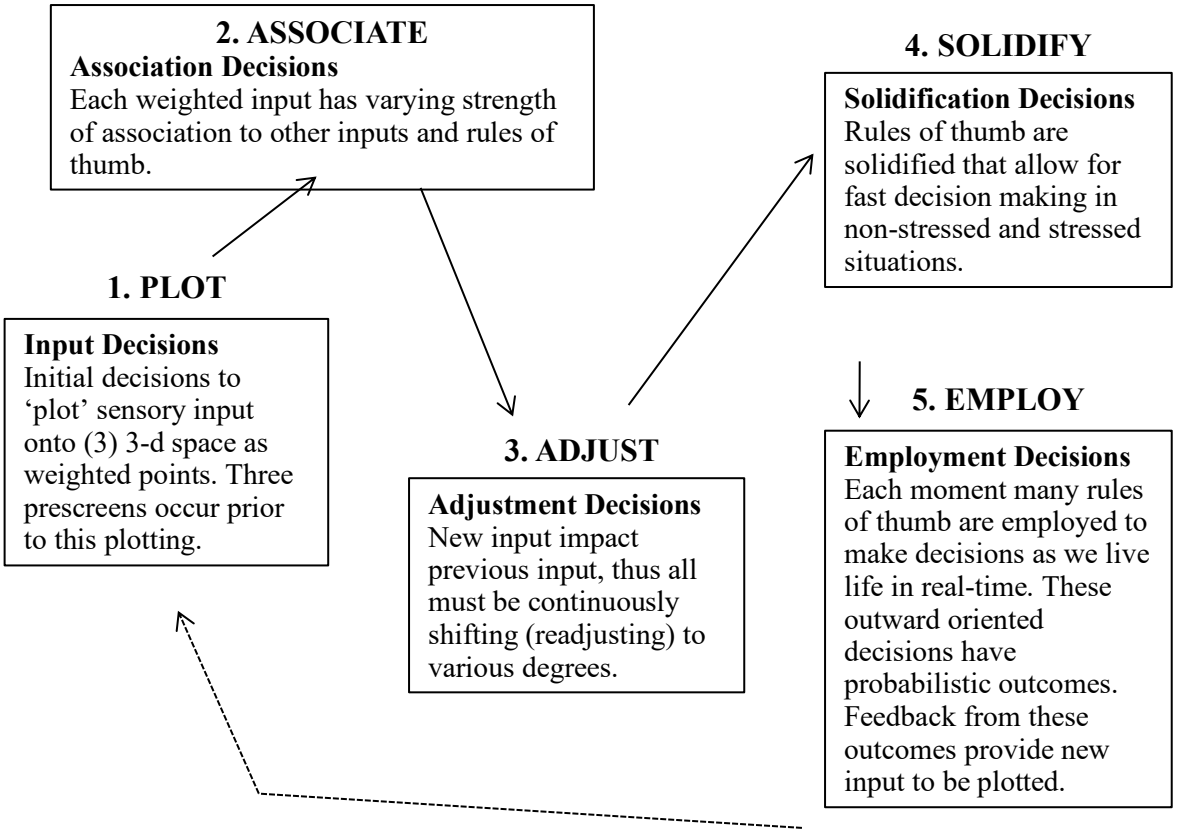


Figure 7: Five Types of Decisions

This section will expand five types of decisions. Each decision (see figure 10) is regulated through the thought dynamo axes.

Plot

Three prescreens occur before input is plotted. First, the input is screened for recognition. Most sensory inputs available to the mind simply "blow through". They are not plotted.

Second, sensory input is prescreened as painful or pleasurable. The immediate basis for such pre-classifications seems to be energy tolerance limits. For instance, if a hot cup of coffee is touched, this prescreen will impact the plotting of this input.

A third prescreening seems to revolve around the notion of comic and non-comic input. This is a more culturally constructed notion that will include the notions of irony, deception and laughter. These notions are really complex and yet input must first go through this prescreen before being plotted. The rules of comic – non-comic are culturally embedded within the (3) 3-D axes.

Pain-pleasure and comic-non-comic continuums serve to modify human thoughts and decisions. For instance, the dog and baby input might be prescreened as comic pleasure – the pleasure of seeing a baby's face in response to the dog's sniffing. Similarly, that same input might be classified by another as a non-comic pleasure – the pleasure of harmony between babies and dogs without any sense of threat. However, this same input could be classified as non-comic pain since the baby was crying and the dog might enhance the threat to the child. And finally a comic pain might result from seeing the pain of a crying baby shift to a response of a bewildered crying tone and a bewildered dog's face as he sniffs. These options are complicated and form a vast array of possibilities within any given situation.

Thus, before input is plotted, these three prescreens are accomplished. Is there recognition of this available input? Is there pain or pleasure from an energy perception? Is there comic or non-comic pain or pleasure within this input?

After these prescreens, the sensory input is plotted for location and weight onto the (3) 3-D axes of logic of intellect, logic of emotion and imagined outcomes (see figure 3). Initial weight of input is directly proportional to emotional and rational intensity of the event.

[Plotting that is essentially feedback from external sources on previous decisions is significantly different from other input in that it links the person with the input. There is an implied or established "I causal" relationship in these inputs.]

Note that words can be plotted in this scheme. All words are spatial at some level of abstraction. All images, music, tastes, smells and touches are also fundamentally spatial – involve particles and waves within physical space. And thus thoughts can be viewed as inputs and are similarly plotted.

Associate

Each input is associated with similar inputs. All input enters through these three center constructs and is disseminated by the strength of attraction from previous inputs organized into rules of thumb. This strength of associations is established through the weights, location and numbers of rules of thumb that have similar spatial/ jealousy/ harmony characteristics.

Strength of association is an attraction factored by previous rules of thumb. Thus, rules of thumb are "seeking" reinforcing input in order to maintain and re-enforce their weight. Lack of new input causes the strength of the rule to fade at some rate. Contradictory input also lessens the weight of the rule.

Adjust

Almost simultaneously, these input weights are modified (through strength of association) by previously plotted inputs and established rules of thumb. New input impacts previous input and previous input impacts new input, thus plotted input must be continuously adjusted to various degrees.

Weights of various types of rules of thumb are directly proportional to number, weights and location of supporting inputs for each rule linked.

Solidify

Solidifying rules of thumb in memories (supported by inputs) is a rehashing of the past in a present. The brain does the hard storage of memories and the mind somehow activates these memories for a "present" rehashing of thoughts, decisions and outcomes. This solidification form and reshape rules of thumb. Within the thought dynamo, rules of thumb are activated.

The TDD model will posit four types of rules of thumb. The first are ***innate rules of thumb***. These are posited as hardwired within human minds across cultures and time periods. These rules are the (3) sets of 3-D axes: logic of intellect, logic of emotions and imagined outcomes. These 9 continuums with 3 central tendencies are deemed innate (i.e. hardwired into the mind); they are apparent from early childhood and form the basis of all other types of rules of thumb.

A second type of rules of thumb involves metaphors. In order to efficiently process large amounts of input, the mind, over time, forms ***image and verbal metaphors***. Tastes, touches and smells are often associated with various words and images. ["It smells like" is a verbal metaphor that is often linked with some image.] Each image and verbal metaphor can be located within the (3) 3-D axes. This simplification speeds the mind to conclusions. For instance, we may have visual and/or verbal metaphor for an older male or female. This type of person may fall within the "father" or "mother" verbal metaphor with many associated thoughts and emotions and imagined outcomes. Similarly an image is usually attached to this metaphor. Thus two people may use the same word metaphor while their image metaphor may be substantially different based upon their previously gathers input concerning "father" or "mother" (Zaltman, 1997, 2000). Verbal and image metaphors constitute a significant agenda for field research.

Third, ***situational rules of thumb*** help us negotiate various circumstances with many real-time factors interacting simultaneously. Situational rules thumb are logical steps of actions when presented with various types of situations. Previously established, these situational rules seamlessly guide much of life. The previously mentioned dog and baby situation is an example. A metaphor rule of thumb for many people may be "precious baby". As the situation unfolds, all input is focused to ascertain one question "Is this precious baby in any threat?" The rule thus implies "I will protect this precious baby if threaten by this dog."

Abstract rules of thumb are a fourth type. An abstraction such as "innocence is precious" is a complex conclusion that can be applied in many situations. These abstractions help mold long-term convictions within people as they negotiate the complexity of life. However, these abstractions, if not thoroughly grounded by innate, verbal and image metaphorical and situational rules of thumb, may simply serve as conceptualization but not as rules of thumb that will govern employment decisions.

More attention is needed to describe abstract rules of thumb. Abstract rules of thumb (see figure 11) are a complex combination of innate, metaphorical and situational rules of thumb. Abstract rules of thumb are higher order rules that shape decision making across complex issues (for instance, the innocence of children is precious or life is valuable). Some people form few abstract rules that they can articulate, while others develop many highly conceptualized abstractions.

Six general abstract questions of reality can account for many abstract rules of thumb. Each of these can be mapped onto (3) 3-d continuums. These abstract rules of thumb form basic convictions/ worldview beliefs of determination (will) that can be employed through making decisions in non-stressed and stressed situations. [Obviously many subsequent questions follow from these six categories – and the categories can be restructured as well.]

1. The Questions of Reality
- Is what we experience real or is it an illusion?
- What is the nature of consciousness? How real are dreams?

2. Foundations of Reality
- What is the nature of matter? What is the nature of energy?
- What is the nature of time and movement? What is the nature of space?
- What is the nature of cause and effect?

3. Authorities of Reality
- What are meaningful meanings? How are meanings internalized?
- What are the meanings of life, work, sex, wealth and recreation?
- What are truth and honesty?
- What are language and communication?
- What is beauty?
- What are intelligence, emotions and imaginations?
- What are the foundational processes of decision making?

4. Relational Realities
- Who am I?
- Does God(s) exist? Who is God?
- What are the natures of humankind, social and cultural relationships?
- What is the self and how is personality arranged?
- What is health on an individual and cultural level?
- Do spirit-beings exist? What is the nature of spirit-beings?
- What is the nature of other life (animal, plant, etc.)?
- How are the young cared for and assimilated into society?

5. Dilemmas of Reality
- What are good and evil? Why is there good and evil?
- What are sin, shame, guilt, and deviant behavior vs. wholeness, peace and joy?
- What are pain and pleasure? Why is there pain and pleasure, beauty and ugliness?
- What are the natures of judgment and mercy?

6. Dependencies in Reality
- What are the basic human needs? What is love?
- What are the natures of life and death?
- How will material wealth be managed in a world of need, greed and beauty?
- What is the drive for human identity?
- What are the purpose and meaning of life?

Figure 8: Various Questions for Abstract Rules of Thumb

Goal Seek

All rules of thumb have goal seeks. That is, these rules seek to achieve a consistency of thought, minimize contradictions, avoid disillusionment and come to a connected peace with paradoxes by negotiating various levels of comic and non-comic pain-pleasure which form stresses.

The primary goal seek of all rules of thumb is a sense of consistency. The mind seeks to be integrated in manageable degrees. Total consistency does not occur, yet a desire for making consistent sense of the world is a continual goal seek.

Within this goal of consistency, the mind seeks to identify contradictions. These contradictions are dealt with by readjusting pervious rules and forming new rules. If no suitable rule is readily available, then a sense of disillusionment is established. This disillusionment may be brief and hardly recognizable or deeply painful and lingering over long periods of time.

If consistency seems impossible and contradiction undeniable, disillusionment may help establish a category of paradox. The establishment of a paradox is a means of resolving contradiction through disillusionment and bringing a new sense of "consistency" – a paradoxical consistency.

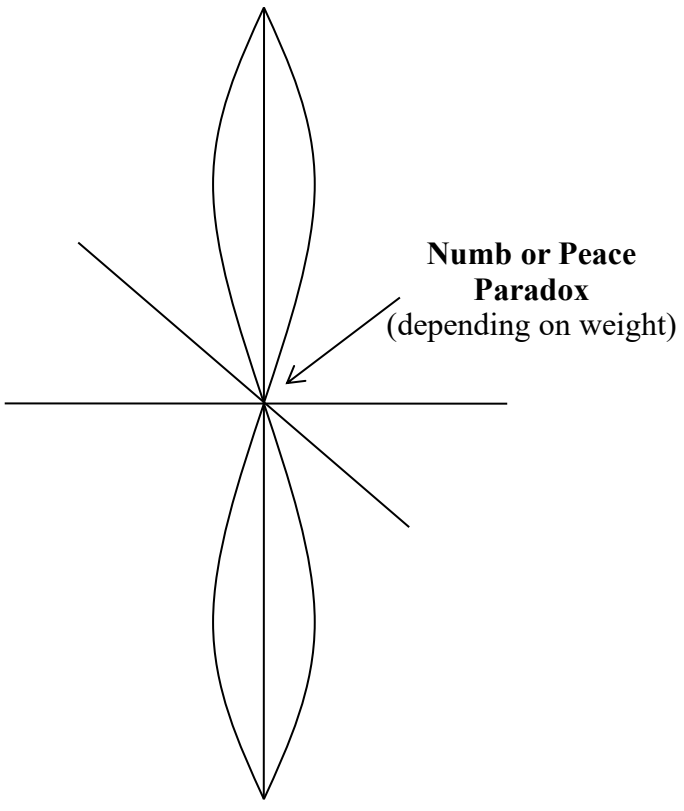


Figure 9: Paradoxes

If proposition "A" and "B" both appear true when considered separately and in conflict when considered jointly, then a paradox has occurred. For instance, the free will of humankind and the sovereignty of God have seemed reasonably true for many when viewed separately. Viewed together they form a paradox.

In this model of decision, paradox is mapped on a vertical axis (see figure 9). This movement to paradox can occur rapidly. What gives the "ah ha" moment of paradox? Sometimes reason and sometimes a more intuitive process involving tapping into the innate category of paradox that can be nurtured over time and with reason. The end-point paradoxes are held with a sense of passion or distain. The central paradox may be acknowledged with an emotional sense of unconnected numbness (ambivalence) or a sense of connected peace; these determinations occur through variations in weight.

A summary of types of rules of thumb and the goal seeks of rules of thumb is given in figure 10.

Types of Rules of Thumb
- Innate Rules of Thumb
- Metaphoric Rules of Thumb
- Situational Rules of Thumb
- Abstract Rules of Thumb

Goal Seeks of Rules of Thumb …
- Seek consistency
- Avoid contradictions
- Avoid disillusionment
- Find paradoxes

These types and goals seeks are in play while negotiating various levels of comic and non-comic pain and pleasure that regulate stresses.

Figure 10: Summary of Rules of Thumb

Employ

All thought decisions (plot, associate, adjust and solidify) are internal. All employment decisions are conceptualized with external consequences of some form. Externally oriented decisions are made by employing the various types of rules of thumb (innate, metaphorical, situational and abstract). Thus speaking out loud, non-verbal expressions and physical movements are viewed as external, i.e. they have direct external interface with the environment.

The thought dynamo is in play as employment decisions are made. That dynamo contains the rules of thumb which make most decisions relatively easy (when compared with the effort that would be necessary to process all relevant data.)

The goals of employment are implicit within imagined outcomes. We imagine outcomes (with varying degrees of lucidity and accuracy) prior to making a decision. Those imaginations of adaptation (thriving – surviving), identity (desired – undesired), and meaning (meaningful – meaningless) are the end goals as we make employment decisions.

After these decisions are made, feedback, usually immediate and often long-term, will be gathered. This feedback is then new input to be plotted, associated, and adjusted and solidified as the future employment of rules are made in real-time.

Under stress, the mind will default to previous rules of thumb rather than attempting to take the time to reason new rules or validate old rules with new input. Later the rules of thumb employed under stress will be associated, adjusted and solidified. There is often a dissonance between non-stressed and stress rules of thumb. (Note: Stress is defined as fear of loss – thus plotted primarily onto the trust-fear continuum.)

All employment decisions contain a logic of yes, no, suspend or create. For instance, "Yes, I decide option A." "No, I decide not to choose option B." "I decide to suspend this decision at this time." "I decided to create new alternatives for this decision set." These four options apply to the imagined outcomes of adaptation, identity and meaning.

Summary of Decisions

In figure 14 each of the five types of decisions are summarized. This flow coupled with the thought dynamo can generate a great complexity of thought and decision.

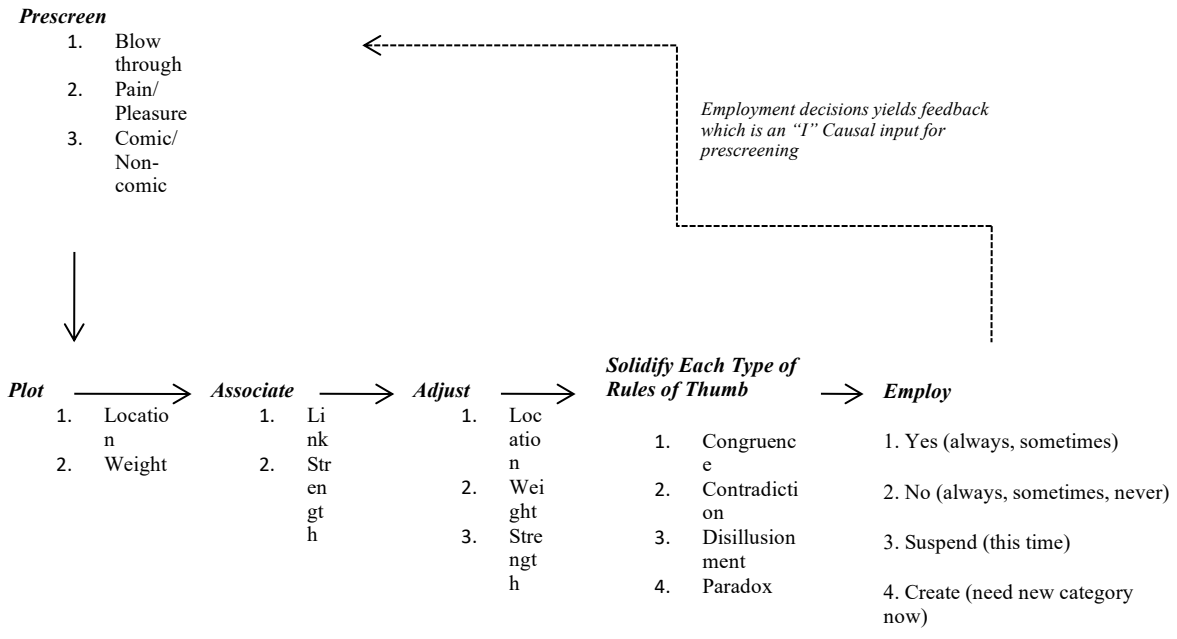
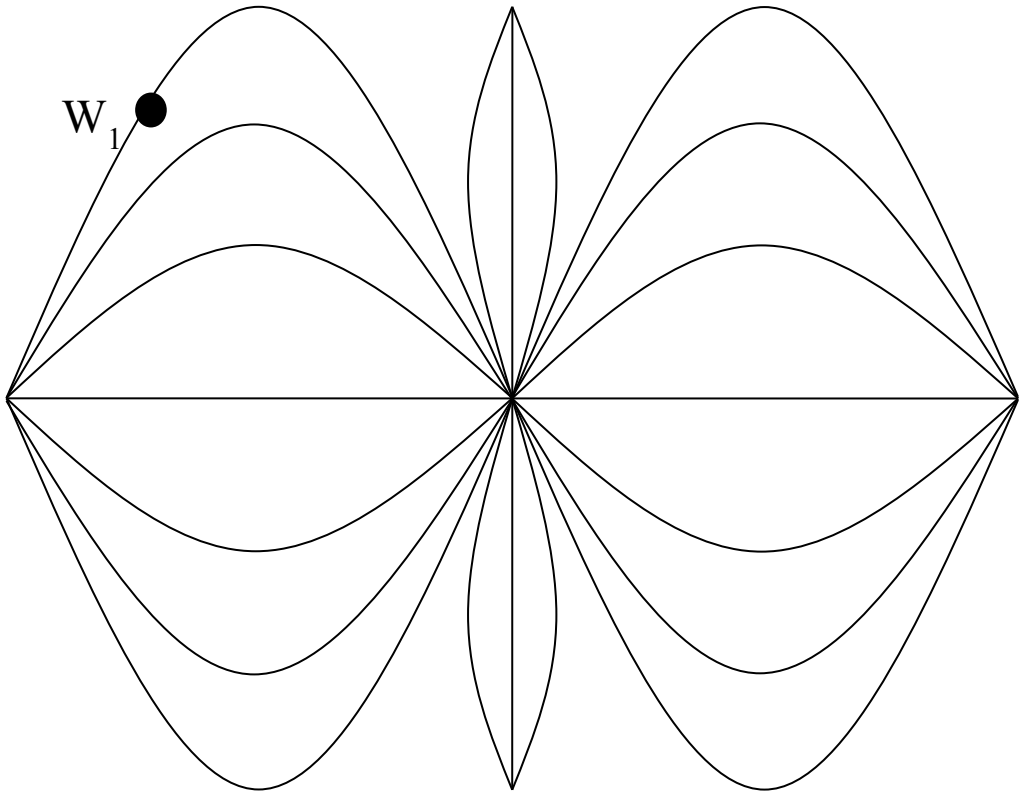


Figure 11: Summary of Decisions

Mapping

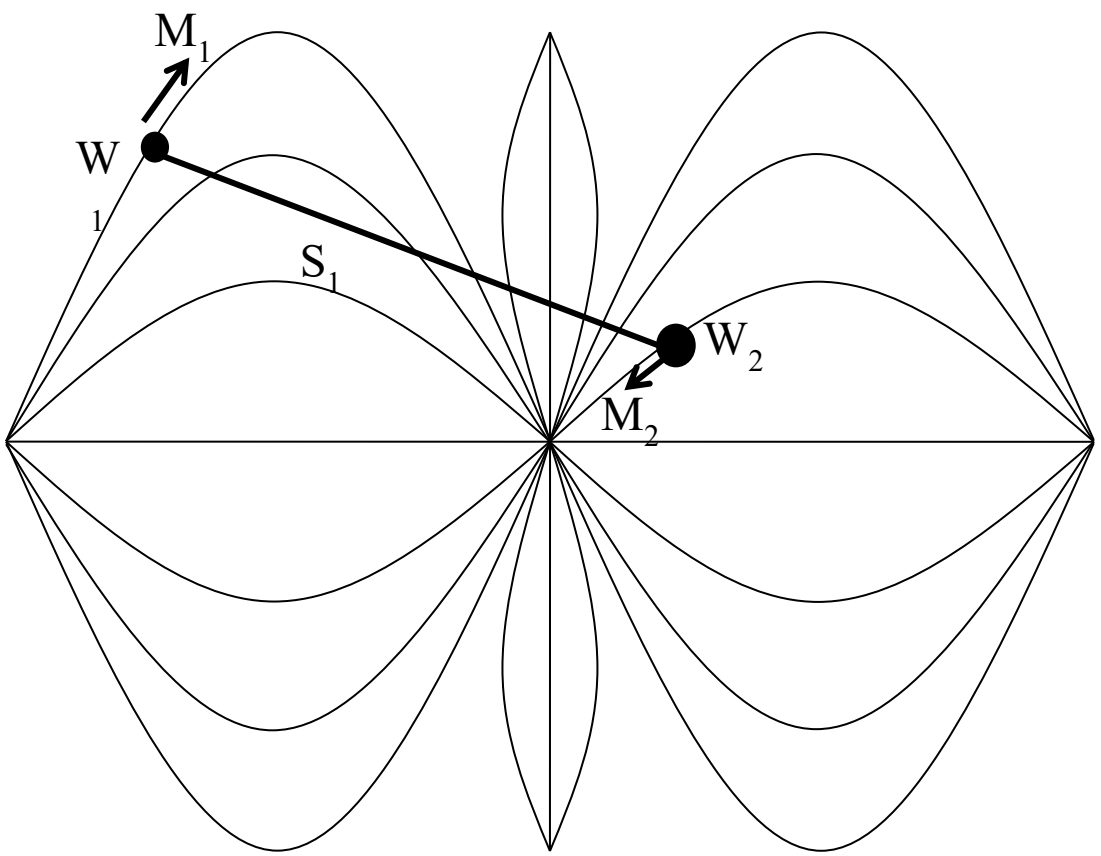
Below are a series of figures (12 – 16) that geometrically describe the mapping of input and rules of thumb within the TDD model.



W – weight of input
Weight is directly proportional to the emotional and rational intensity of the input event

Figure 12: Plot Input

In figure 12 an input has been assigned a location on the (3) 3-D axes and a weight.

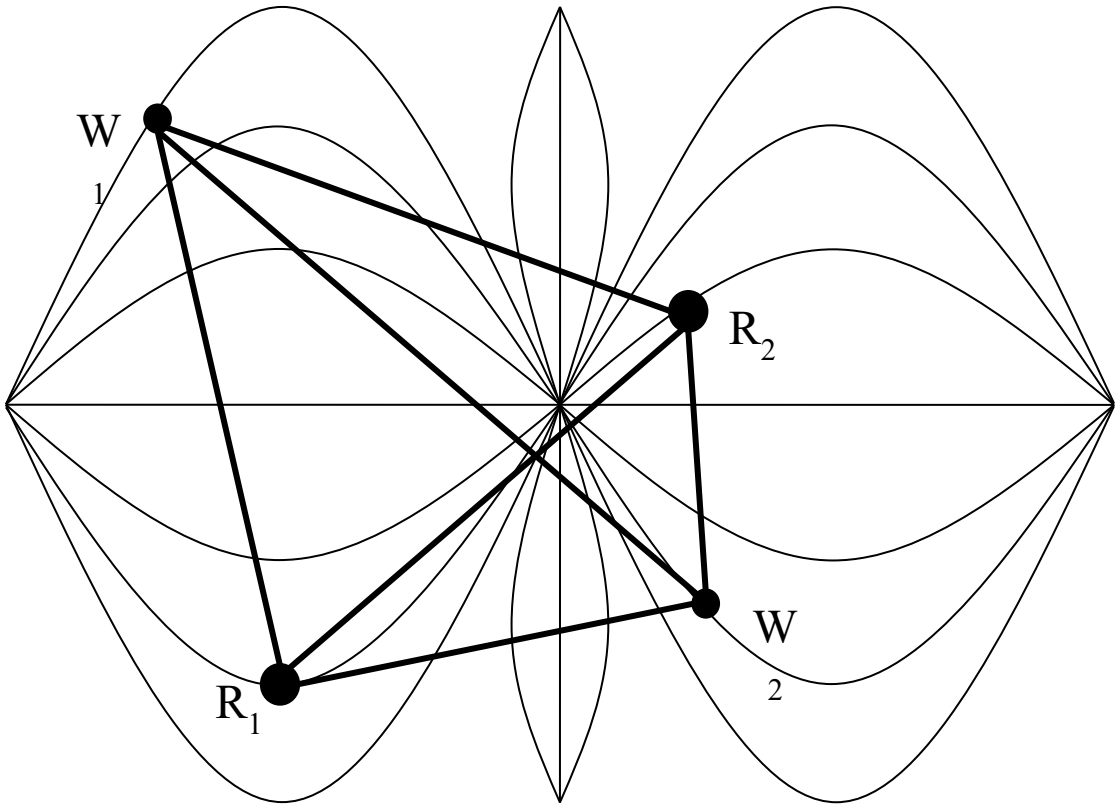


W – weight of input, S – strength of association (as in spring tension), M – adjustment movement, Weights can jump up or down due to the strength of other weights or rules.

Figure 13: Associate and Adjust

In figure 13 an association with previous input has been established and the strength of that association has been assigned. A subsequent adjustment is made as these two inputs impact each other. There is a similar process for weights and rules of thumb.

If this association exceeds some strength, then plotted input can 'jump' the sine rails. Also in time the strength of association can weaken or strengthen.



The inter-linking aspect of the model may involve large numbers of points.
W – weight of input; R – rule of thumb

Figure 14: Solidify Rules of Thumb

In figure 14 weighted inputs are impacted and solidified by rules of thumb and vice versa.

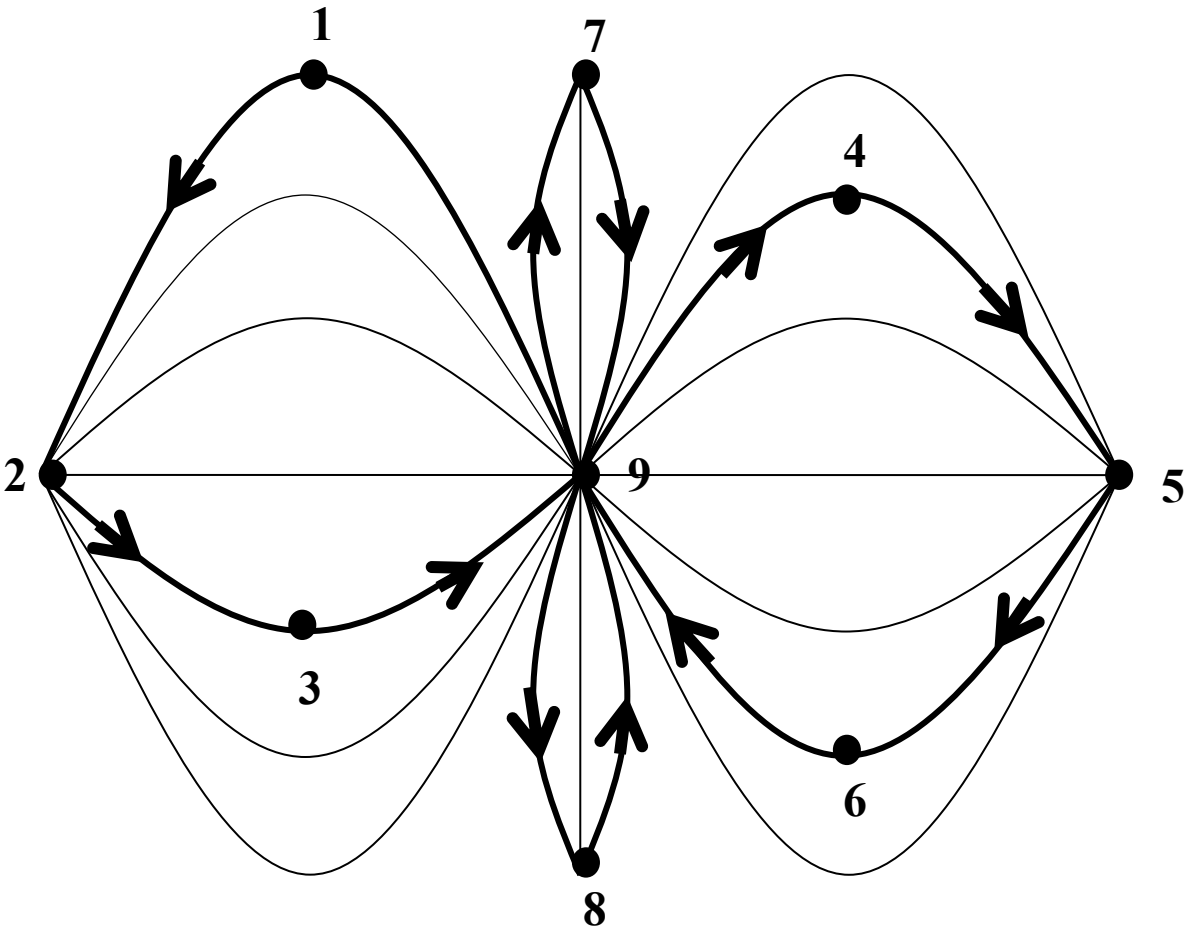


Figure 15: An Example of Geometric Pathways

An example of pathways for input and rules of thumb is given in figure 15. This pathway is one of many that can be conceptualized.

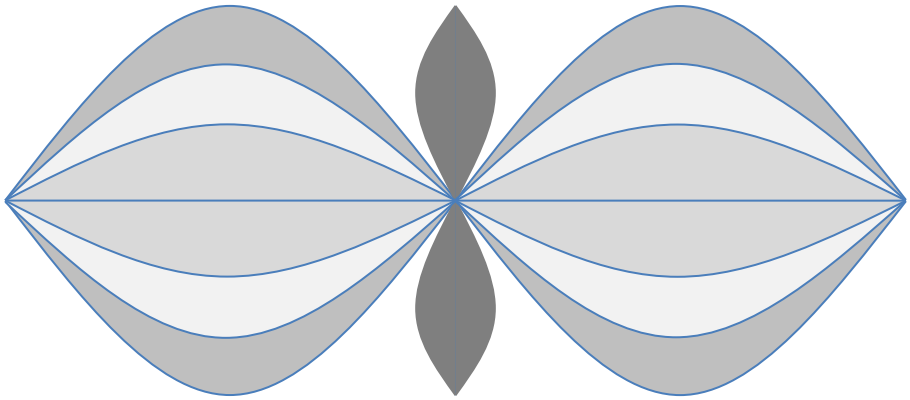


Figure 16: (3) 3-D Axis Overlay Dynamics

As employment decisions are made, these decisions become new "thought" inputs and the cycle of plotting, association, adjusting solidifying and employing continues and are mapped accordingly.

It should be noted that all these plots, associations, adjustments, solidifications and employment decisions occur with the (3) 3-D aspect of this model (figure 16). The central tendency of this overlap is creative harmony of jealous space.

Rules for Optimization

This section will suggest logical rules for optimization of decision making under multiple weighted points over varying periods of time. The primary assumption in optimization of decision making is the selection of one or more points to be optimized within the weighted spherical modeling of decision making. In the case of a singular point, optimization is geometrically apparent – consistency in decision making to achieve congruence that increases or maintains the weight of a particular locus point. For instance, if a particular desired identity is a singular point, then all decision making can be focused on achieving that optimal identity. However, most decision making is far more complicated. Multiple points for optimization are the norm, not the exception. Achieving a desired identity at the cost of good-evil or fear-trust is not optimal decision making for most contexts.

In order to model the optimization of multiple points with varying intensities, the concept of stress lines between points must be managed through decision making. These lines can be lines of congruence, contradiction, disillusionment, or paradox. In other words, these four goal-seeks of decision are key to the process of optimization of multiple points. Within the spherical model, proximity of weighted points can be managed for optimization by the reduction of intensity of stress line between two or more points. The intensity of a stress line is directly proportional to: 1) the weight of each locus point, 2) the distance between loci (the greater the distance the greater the

stress up to some point and then a process of detachment begins), 3) the distance from the central construct of the interrelated axes to the various points, and 4) the stress of indecision over time. The goal in optimization is to manage stress lines in a way that forms a paradoxical central points (not a congruent point, a contradictory point, or a disillusionment point). This paradoxical central locus point is a fuzzy point at or near the central construct of the interrelated axes. The greater the distance from the central construct of the interrelated axes, the more likelihood that the goal of long-term optimization will degenerate into congruence for short-term optimization or contradiction (failure to achieve an optimal decision) or disillusionment (a decision to abort an optimization process).

Thus, the fuzzy optimal decision of multiple loci in multiple contexts over a long-term can be modeled as a paradoxical solution in the vicinity of the central construct of the interrelated axes. The optimization will remain fuzzy in that it is viewed as a 'fluid particle' solution rather than a solid plotted point within the weighted sphere.

Mapping the Japanese Construct 'Amae'

In order to establish the utility of the TDD model across cultures, the Japanese construct of *'amae'* will be considered. Below (see figure 20) is a means of mapping *'amae'* onto the (3) 3-D axes of the thought dynamo. This example is chosen to address this question: Can the TDD model reasonable address intellectual and emotional reasoning and imagination across cultural boundaries?

> Wierzbicka (1993) does not perceive understanding emotions as an easy endeavor, but rather one that requires difficult translation. She stated, "I maintain, however, that no matter how 'unique' and 'untranslatable' an emotion term is, it can be translated on the level of semantic explication in a natural semantic metalanguage and that explications of this kind make possible that 'translation of emotional worlds' (Lutz 1985a) which seems otherwise impossible to achieve" (p. 135). I view the translation of complex culturally constructed emotions as a crucial issue. The TDD model must be able to aid in emotional translation in order to be considered a culture-general model of thought and decision.
>
> Japanese psychiatrist Takeo Doi (1981) described in detail the dynamics of *amae* in the Japanese culture stating, "The Japanese term *amae* refers, initially, to the feelings that all normal infants at the breast harbor toward the mother – dependence, the desire to be passively loved, unwillingness to be separated from the warm mother-child circle and cast into a world of objective 'reality' " (p. 7). He went on to say, "… all the many Japanese words dealing with human relations reflect some aspect of the *amae* mentality. This does not mean, of course, that the average man is clearly aware of *amae* as the central emotion in *ninjo* (human feeling)" (p. 33). Furthermore, he compared the Japanese with Westerners in stating, "Scholars have put forward many different theories concerning the ways of thinking of the Japanese, but most agree in the long run that, compared with thought in the West, it is not logical but intuitive" (p.76). Doi proposed outsiders struggle with the *amae* construct. He stated, "… to persons on the outside who do not appreciate *amae,* the conformity imposed by the world of *amae* is intolerable, so that it seems exclusivist and private, or even egocentric" (p. 77). Thus, *amae* serves as an appropriate test for the culture-general claim of the TDD model.

The TDD model can be used to translate the Japanese construct of *amae*. This meta-language construct is not intended to fulfill the richness of the Japanese construct but rather to approximate its construction in such a way that translates *amae* into the meta-language of the TDD model. The *amae* construct is the dependency implied in leaning on the goodwill of benevolent other(s). It requires **trust** in other(s). It implies the **fear** of being betrayed by others. It requires the **bonding** of dependency. It yields the **freedom** of dependency. It requires the **honor** of submitting to another's will. It forbids the **shame** of betraying another. It requires the management of a privileged and thereby **jealous** relationship between people. It yields the **power** of being provided for. It requires the **powerlessness** of receiving. It requires **intuition** to negotiate relationships. It assumes the **accurate** interpretation of *amae* as a social construct. It requires an acknowledgement of **good** in one's in-group. It hold that **evil** is betrayal of one's in-group. It requires the negotiation of **space** between two or more people. It requires the proper networking of relationships for both **surviving** and **thriving**. It requires a **desired identity** of being fundamentally a self that is dependent. It views the absence of a dependent relationship as an **undesired identity**. It views the parent-child relationship as the fundamental **meaningful** relationship. It views the absence of *amae* as fundamentally a **meaningless** existence. It requires both persons in the relationship maintain and **creatively** enhance **harmony**.

Thought Dynamo	**Japanese "*Amae*"**
Logic of Intellect	
Powerful – powerless	*Amae* requires the powerlessness of receiving and yields the power of being provided for.
Good – evil	*Amae* requires an acknowledgement of good in one's in-group and holds that evil is betrayal of one's in-group.
Accuracy – intuition	*Amae* requires intuition to negotiate relationships and assumes the accurate interpretation of *amae* as a social construct.
Space	*Amae* requires the negotiation of space between two or more people.
Logic of Emotion	
Trust – fear	*Amae* requires trust in other(s) and it implies the fear of being betrayed by others.
Honor – shame	*Amae* requires the honor of submitting to another's will and it forbids the shame of betraying another.
Freedom – bonding	*Amae* requires the bonding of dependency and yields the freedom of dependency.
Jealousy	*Amae* requires the management of a privileged and thereby jealous relationship between people.
Imagined Outcomes	
Surviving – thriving	*Amae* views the proper networking of relationships for both surviving and thriving.
Desired identity – undesired identity	*Amae* views self as dependent as a desired identity and views the absence of a dependent relationship as an undesired identity.

Meaningful – meaningless	*Amae* views the parent-child relationship as the fundamental meaningful relationship and the absence of *amae* as fundamentally a meaningless existence.
Creative harmony	*Amae* requires both persons in an *amae* relationship maintain and creatively enhance harmony

Figure 17 A Translation of *Amae*

Computer Modeling

The geometrical nature of the model lends itself to mathematical equations and computer modeling. If input can be plotted, associated, and adjusted and the dynamics of rules of thumb calculated, then decisions might be predicted or reasonably formulated using sophisticated computer modeling. By accurate predictions, the model can be assumed to be reasonably sound.

The logical rules of conventional computer programming fail to mimic human intelligence. The common sense needed to address child like questions also escapes current AI models (Minsky, 2007). The TDD model can be used in artificial intelligence programming to better approximate human intelligence. The TDD model accounts for the nuance of emotions as well as such issues as paradoxical reasoning and comic repositioning. Though not developed here, the TDD model can be a pathway for future AI advances and the full utility of the TDD model can be realized within AI computing.

Future Research

The TDD Model is appropriate for descriptive, prescriptive and predictive research.
Regarding descriptive research, inputs can be gathered through images as well as word.

For instance participants can be asked to rate images on a line scale for statements such as "This image represents power to me" or "This image engenders trust for me" or "This image invokes a sense of meaninglessness to me." Some images may have singular subjects; other may have a subject in a "usual cultural context" while another may have one primary subject in an "unusual cultural context". And finally many interrelated subjects in 'usual' and 'unusual' cultural contexts may be rated. Many people are required for valid conclusions, thus each person may be asked address only one or two continuum per image. Much input can come through the sorting and selection process of images on Internet search engines such as Google.

Inputs of words are similar to the above. One option within the publishing industry is to ask demographic and the (3) 3-D continuum questions for book titles and covers as they are sold.

In addition a decision preference inventory using a Likert scale has been developed to help identify individual and group decision preferences.

Conclusion

The dynamo of thought drives the dynamics of decision. The TDD model has posited an interaction between the logic of intellect, logic of emotion and imagined outcomes. This innate interaction of the mind is in play as five types of decisions are made: plot decisions, association decisions, adjustment decisions, solidification decisions and employment decisions. This model can serve toward descriptive, prescriptive (optimized) and predictive research objectives.

References

Buss, D. (2000). The dangerous passion: Why jealousy is as necessary as love and sex. New York, NY: The Free Press.

Byrnes, J. (1998). The nature and development of decision making: A self-regulation model. Mahwah, NJ: Lawrence Erlbaum Associates.

Doi, T. (1981). The anatomy of dependency: The key analysis of Japanese behavior. Tokyo, Japan: Kodansha International.

Ennis, Ralph (2004). "A Theoretical Model for Research in Intercultural Decision Making", *Intercultural Communication Studies, Volume XIII*, pp 113-124.

Fishburn, P. C. (1970). Utility theory for decision making. New York: John Wiley and Sons.

Hart, S. and Carrington, H. (2002) "Jealousy in 6 month old infants", *Infancy, Volume 3*, pp 395-402.

Kulich, S., Ennis, R. & Rhey, B. (2001). Cross-cultural dimensions of decision-making motives: Analysis in a Chinese framework with applications to business in China. Paper presented at The American Association of Behavioral & Social Sciences, Las Vegas, Nevada, 2001.
Kuula, M. & Salminen, P. (1996). Prospect theory and mood. In P. Walden, M. Brännback, B. Back & H. Vanharanta (Eds.), The art and science of decision-making in honour of Christen Carlsson (pp. 127-137). Painosalama Oy, Åbo: Åbo Akademi University Press.

Lutz, C. (1983). Parental goals, ethnopsychology, and the development of emotional meaning. Ethos, 11, 246-262.

Maruyama, M. (1987). Communication between mindscape types. In J. VanGigch, (Ed.), Decision making about decision making: Metamodels and metasystems (pp. 79-97). Cambridge, MA: Abacus Press.

Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. Genetic, Social, and General Psychology Monographs, 12 (3), 339-361.

Minsky, M. (2007). "The Discover Interview", *Discover, Volume 1,* pp.14-18.

Poulson, C. (2000). "Shame: The Master Emotion?" University of Tasmania, School of Management Working Papers Series.

Tsien, Joe Z. "The Memory Code", *Scientific America, July 2007*, pp. 52-59

Tversky, A. & Kahneman P. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

U. S. Project. (1997). Louisville, KY: Cultural Insights, Inc.

Wierzbicka, A. (1992). Semantics, culture, and cognition: Universal human concepts in culture-specific configurations. Oxford, England: Oxford University Press.

Wittgenstein, L. (1958). Philosophical investigations (2nd ed.). (G. E. M. Anscombe, Trans.). Malden, MA: Blackwell Publishers. (Original work published 1953.)

Zaltman, G. (1997). Rethinking market research: Putting people back in. *Journal of Marketing Research, Volume 34*, pp 424-437.

Zaltman, G. (2000). Consumer researchers: Take a hike! *Journal of Consumer Research, Volume 26*, pp 423-428.

A Theoretical Model for Research in Intercultural Decision Making*

**This information is also included in Wharf and Bearing VIII.*

Below is the reference for my brief article on intercultural decision making.

> Ennis, Ralph (2004). "A Theoretical Model for Research in Intercultural Decision Making", *Intercultural Communication Studies, Volume XIII*, pp 113-124.

And the table of contents for my thesis as required for a Masters in Intercultural Relations is below:

TITLE: A Theoretical Model for Research in Intercultural Decision Making

CHAPTER 1. INTRODUCTION **1**

CHPATER 2. METHODOLOGY **6**
- Methodology of Theory Construction
- Qualifying the ICDM Model
- Using Conceptual Space in Model Development
- Philosophical Assumptions of the ICDM Model
- Rationale for Constructing the ICDM Model

CHAPTER 3. LITERATURE REVIEW **15**
- Decision Making Literature
- Expected Utility Model
- Constraints Model
- Expectancy-Value Theory
- Health Belief Model
- Prospect Theory
- Social Judgment Theory
- Self-Regulation Model
- Reciprocal Causal Models
- Interdependency and the Negation of Super-Logic
- Summary of Decision Making Literature
- Consumer Behavior Literature
- Engel-Kollat-Blackwell Model
- Howard-Sheth Model
- Nicosia Model
- Mullen and Johnson's General Model of Consumer Behavior
- Fishbein and Ajzen's Model of Belief, Attitude, Intention and Behavior
- Assael's Model of Complex Decision Making
- PAD Emotional-State Model
- ZMET Research Process
- Cultural Insights, Inc. Process
- Summary of Consumer Behavior Literature

Intercultural Literature
Anxiety/Uncertainty Management
Violation of Expectancy Values
Identity Negotiation Theory
Cross-Cultural Adaptation
Shared Meanings
Systems Theory
Hall's Low and High Context Communication
Bennett's Developmental Model of Intercultural Sensitivity
Erickson's Life Cycle Model
Moral Development Models
Learning Models
Summary of Intercultural Literature
A Composite of Factors from Literature

CHAPTER 4. AN INTERCULTURAL CONTEXT FOR DECISION MAKING 47
Defining the Intercultural Context
The Decision Maker(s)
Primary Nature of Decision Making
Overall Flow of Decision Making
Inputs
Interactive Decisions
Internal Decisions
Culturally motivated sensory perception decisions
Self and others assessment decisions
Belief integration decisions
Emotional Integration Decisions
Change Assessment Decisions
Behavioral Communicative Decisions
Outside Influence Factors
Outcomes of Decisions
Feedback from Outcomes of Decisions
Cultural Change through Decision Making
Importance of Intercultural Decision Making
Chapter Summary

CHAPTER 5. A MODEL OF INTERCULTURAL DECISION MAKING 89
ICDM Model – A Process Model of Intercultural Decision Making
Logic of Intellect
Power Axis: A Continuum from Powerful to Powerless
Morality Axis: A Continuum from Good to Evil
Certainty Axis: A Continuum from Accuracy to Intuition
Central Construct of Logic of Intellect: Space
Logic of Emotion
Liberty Axis: A Continuum from Freedom to Bonding
Relational Axis: A Continuum from Trust to Fear
Hierarchical Axis: A Continuum from Honor to Shame
Central Construct of Logic of Emotion: Jealousy
Imagined Outcomes
Adaptation Axis: A Continuum from Surviving to Thriving

Identity Axis: A Continuum from Desired to Undesired Identity
Meanings Axis: A Continuum from Meaningful to Meaningless
Central Construct of Imagined Outcomes: Creative Harmony
Minimal and High Stressed Decision Making
An Intersecting Dynamic of Decision Making
A Special Case: The Japanese Concept of Amae
Scope of the ICDM Model
A Predictive Model
A Developmental Model
Ideas for Future Research
Chapter Summary

CHAPTER 6. APPLICABILITY OF THE ICDM MODEL IN RESEARCH 137
An Example: Chinese and American Intercultural Decision Making
Logic of Intellect
Logic of Emotion
Imagined Outcomes
Intersecting Concepts
A Tentative Comparison of American and Chinese Decision Making
Some Implications for American-Chinese Intercultural Relations
Chapter Summary

REFERENCE LIST 160

APPENDIX 168
Research of American and Chinese Students

BIBLIOGRAPHY 172
References from Decision Making Literature
References from Consumer Behavior Literature
References from Cultural and Intercultural Literature
References regarding Chinese and Americans

Resources to Explore

CONNECT BIBLE STUDIES

- GOD: Connecting with His Outrageous Love
- IDENTITY: Becoming Who God Says I Am
- SOUL: Embracing My Sexuality and Emotions
- RELATIONSHIPS: Bringing Jesus into My World
- LIFE: Thriving a Complex World
- FREEDOM: Breaking the Power of Shame

IMAGE-BASED

- The New Me
- Searching the Ordinary for Meanings
- Grapplings: Why Do People Suffer So Much?

BOOKS

- The Shame Exchange: Trading Shame for God's Mercy and Freedom
- Worth a Thousand Words: The Power of Images to Transform Hearts

INVENTORIES

- Breakthru: Discovering My Spiritual Gifts
- Breakthru: Discovering My Primary Roles

For the above resources see:

www.ralphennis.com

About the Authors

Ralph and Jennifer Ennis have served with The Navigators since 1975. They have ministered at Princeton University, Richmond Community, Glen Eyrie Leadership Development Institute, The CoMission in Russia, and in Raleigh, NC. In 2006 Jennifer co-founded JourneyMates, a ministry to help people grow in intimacy with the Triune God through Scripture, silence and solitude.

Unless otherwise noted, the essays of the WB Series have been written by Ralph. However, each work was crafted in the context our marriage relationship and with the editorial benefit of Jen's perspectives and unique abilities.

In 2018 Ralph and Jennifer celebrated 45 years of marriage. They have four married children and 15 grandchildren.

Our web sites:

www.journeymates.org
www.ralphennis.com

Cover and About the Authors updated December 2018