A Brief Explanation of Consciousness

M. E. Tson

This short paper demonstrates how subjective experience, language, and consciousness can be explained in terms of abilities we share with the simplest of creatures, specifically the ability to detect, react to, and associate various aspects of the world.

Instinct. With the innate ability to *detect* (an event or reaction that, like rust, only takes place in the presence of certain conditions), *react* to (accept, reject, destroy, reproduce, explore, flee, etc.)[1,2,3] and *associate* (to form/strengthen via experience new connections among detections and reactions so that the activation of one makes the activation of the other more likely)[4] certain internal and external stimuli, an organism can begin to form *de facto* categories (some things are food or a mate; some aren't),[5] develop expectations, and, to some extent, have its instinctual reactions modified by experience. Although the organism's actions and decisions can still be completely explained in terms of instinct - it can learn and problem solve. We can see such capabilities in numerous other animals, but there is no need to assume that they have any self-awareness or are anything more than *a contiguous collection of chemical reactions, which happen, by evolutionary 'selection', to all work in a coordinated way to further "their" collective survival.*

Communication. Many young animals learn various survival strategies (hunting, fishing, social habits, means of communication, etc.) from their community. Macaques, vervet monkeys, cockerels, prairie dogs, lemurs, Californian ground squirrels, and chickens have all been found to learn and use differentiated alarm calls or "vocabulary" to indicate different threats.[6,7,8] Just as a young vervet learns to climb down when it spies an eagle and to climb up when it spots a leopard, it also learns that one shriek is emitted in the one instance and a different sound in the other.[9] Communication learning involves associations about *when to say* as much as *what to say*. *Rather than thinking of communication as sounds associated with objects, it is better to think of them as being associated with experiences (which can include objects).*[10] As with the senses, the communication processing areas of the brain seem specially wired to find patterns and form associations between contextual (particularly social) circumstances, vocal sounds, and their order. One difference, in degree if not in kind, in the human case is our ability to articulate thousands of different sounds[11] and to form unique associations with each of them.

Like the young macaque, a child learning the word "flower" or big sister "Angela" is not learning their definitions, but when (in what situations) the words "flower" or "Angela" are used. There is no single place in the brain where a word is defined.[12] Words like "Angela" and "flower" are simply associated with the numerous other associations and expectations about their properties and tendencies which experience has formed: with "flower" what happens when you look at it, smell it, pull it, etc., and a comparable but much more complex set of expectations for "Angela." If we add the word *look*, as in "*Angela looks at the flower,"* certain of these associations, and those regarding one of Angela's properties, in particular (expectations regarding eye gaze and another's attention and intentions), are triggered. "*Angela picked the flower"* triggers slightly different associations. A given sound's several possible meanings are usually easily distinguishable, because the associations triggered by hearing, "I'm going to the *bank*" will differ depending upon whether the speaker is holding a checkbook or a fishing rod.[13]

Learned words are sounds that are associated with--and thus serve as additional mental cues for--other associations experience has formed. When I meet another English speaker and use a word like "chair", "mother," or "honor", I'm assuming that she has had experiences that are similar enough to mine so that she has formed similar associations with that word. *A language*

isn't shared words and syntax, it is shared or, rather, convergent, experiences which are associated with those words/syntax and which give them their meaning.[14] As previous linguistic associations are applied to unprecedented experiences, individuals are capable of uttering intelligible phrases which they themselves may have never heard before.[15]

Our capacity to form associations is wider and deeper in some areas than other animals but not substantially different. Rudimentary displaced reference and abstract thought can be seen in the fact that it is not necessary for the other macaques to actually see the approaching threat themselves before they respond to the warning by scampering in the appropriate direction for safety.[16] The greater the number of successive associations which separate a word from detections and innate reactions, the more 'abstract' it is, but this is no less mechanical than anything we have discussed up to now. None of this requires that either party in a communication be aware of the fact that it is communicating. One instinctual or learned reaction simply follows another.

Self-awareness. Detection, reaction, categorization and associations, problem-solving, community, … everything up to and including communication could have existed long before there was any creature who could observe, label, and comment on itself or the world around it. Language provides the external frame of reference necessary for labeling and reflection. A phrase like, "Angela picked the flower" leads to a certain set of mental associations and expectations about Angela and the flower, and these associations are themselves made up of other, more basic ones. Leaving aside the different subjective reaction produced in the speaker, there is no cognitive difference between saying, "She picked the flower." or "I picked the flower." The self is subject to observation, description, evaluation, and labeling just like any other object and reflection thus becomes possible. We observe and label ourselves using the same linguistic associations that we use to describe other objects and experiences. As can be seen in experiments with split-brain patients, beyond our capacities of detection, reaction, and association, estimations as to our own inner, unseen motivations are just as speculative as those of any other observer.[17] 'Just as the young monkey can learn when to emit the "flying predator" warning,[18] a child learns in which circumstances to utter "I" or "me"; just as a young cub can learn what to do when it is hungry, a human infant can learn what to say when it is hungry. But once the infant has learned in which situations to say "*I* am …", it enters a whole new world of possibilities in self-awareness and subjective experience.

Although we often equate inner speech with *thought*, *the mind* and *consciousness*, it is only the "tip of the iceberg" of the reflexes, reactions, and associations that we have been talking about up until now. Nor is it necessarily or automatically the executive center of decision-making. Through inner speech, we can reflect upon our condition and--by bringing additional associations to bear--influence the focus of our energies and efforts.[19] Nevertheless, while the conclusions we reach--like 'I should stop smoking' or 'I am a coward'--introduce additional associations and reactions into the decision-making process, they are not the final word. Learned associations do not erase innate patterns of detection and reaction and may even conflict with them. The result is a *Man vs. Himself* struggle.

Personal identity has a continuity through time because of social conventions, roles, and the fact that our *past* associations or memories determine our *present* outlook on how best to approach our *future*.[20,21]

Metacognition. In short, we are not aware of what we are thinking. We are aware of what we have just thought.[22] Only when we stop to reflect and remember do these past thoughts and actions then become the objects of present cognition. When someone unexpectedly asks, "What color are you looking at?" we cannot answer them instantly, but we must stop and search our immediate short-term memory, which we have learned to use as a proxy for the present.

Metacognition, which uses short-term memory and inner speech, is always at least one step behind cognition. We think we are aware of our present selves, but we are really only self-aware of our immediate past selves.

The Hard Problem: Qualia. An organism's subjective experience has several components. ***First***, detection. Because unique nerve combinations fire, different sensations are immediately distinguishable. When we see the color 'orange' what we are as importantly seeing is something that is not red, yellow, blue, purple… (Recalled images involve reactivating the same "early sensory" neurons triggered by external stimuli.)[23] This component alone is no more impressive than a camera. ***Second***, reaction or changes in relative attention and energy allocation.[24] Although culture will help to further define and extend upon what is seen as good or desirable, and self-awareness may eventually cause us to question the propriety of our innate drives, all urges and actions have their origin in one or more of the innate primary reactions. All meaning and purpose that we see in the world arise from these instinctual reactions. We cannot conceive of or experience the world outside of them. They are our world. *Usually, that world is some complex, interacting mixture of several reactions to stimuli,* but at the highest levels of intensity (rage, terror, ecstasy), it becomes one instinctive reaction or inclination in particular which becomes the focus of all our energies and attention.[25] We exist only to satisfy that drive; nothing has importance outside of the satisfaction of *that* drive. Thus, we can understand and explain what it feels like to experience 100% intensity of a particular basic emotion. (At the highest levels of intensity--rage, terror, ecstasy--, our world becomes one instinctive reaction or inclination in particular which becomes the focus of all our energies and attention.[26] We exist only to satisfy that drive; nothing has importance outside of the satisfaction of *that* drive.)

Every other emotional state (and these are innumerable[27]) is a mixture of less intense instances of these primary reactions.[28,29,30] But the feeling of 75.6% of one drive plus 20.4% of another plus 4% of a third is a bit more difficult to express. The situation is analogous to that of a painter who has production formulae and names for the three primary paints but who does not have any labels or precise formulae for the 10 million[31] other shades she can also produce by mixing these three. Although the painter lacks and may never have specific names, or precise formulae to produce, each of the millions of other possible shades, it would be a mistake to conclude that this **'explanatory gap'** (or better put 'descriptive gap') meant that the as yet undefined shades were therefore mysterious, unexplainable or irreproducible from the primary three for which she does have precise words and formulae.

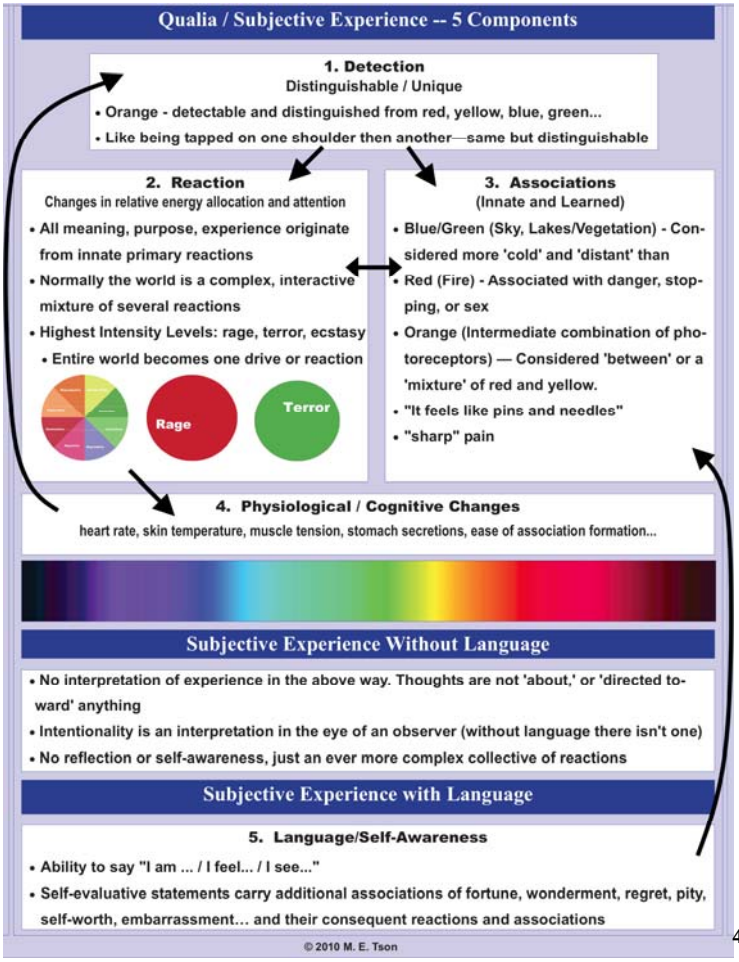
Similarly, although we do not have descriptions for the innumerable, immensely complicated emotional states that arise from lower intensity combinations of the basic reactions whose high-intensity feelings we can describe, this should not be taken to mean that our un-described feelings must result from, or contain, some heretofore-unexplained phenomenon. So the *second* aspect of understanding subjective experience is what portion of an organism's attention or energies is directed towards and invested in particular drives.[32]

Third, innate and learned associations. Green and blue (the colors of vegetation, lakes, and sky) are considered more 'cold' and 'distant' (short-wave light and distant objects require similar optic adjustments) than, say, red (the color of fire), which in different contexts may be associated with danger, stopping, or sex.[33] We may associate orange as being 'between' or a mixture of red and yellow as an intermediate combination of photoreceptors fires.[34] Blue may be associated with "darkness" more than, say, yellow because the eye is less sensitive to the lower wavelengths.[35] Subjective experience is also always in some context. A "colored" object may be edible, climbable, a predator, wood, shiny, or rough. These additional associations and their consequent reactions are combined with whatever associations and reactions the color itself gives rise to. Distinct associations (and detection) account for the difference between being frightened

(same reaction) by a 'snake or a bank manager.'[36] ***Fourth,*** accompanying physiological and cognitive changes. The neurotransmitters that help to give rise to reactions also affect the speed of image formation and change.[37] During the reactions of acceptance and reproduction (happiness), associations are formed more rapidly and are not held for as long. Association is freer and may even become over-inclusive. Motor efficiency and exploratory behavior are increased while inhibition is decreased.[38] Additionally**,** our experience of a reaction like fleeing or protection isn't merely the preparation to fight or flee. It is also to detect our heart quicken and our muscles tighten.[39] We detect many of these and other internal changes just as we detect external stimuli, and we have reactions to and associations with them, as well.[40]

The above four components of subjective experience are shared with other animals[41] although, without language, there is no interpretation of experience in this way. Thoughts are not yet 'about,' or 'directed toward,' anything. *Intentionality is an interpretation in the eye of an observer*[42] *and without language (to explicitly separate the train of reactions into subjects and objects) there isn't one.*[43] At this stage, there is still no need to assume that there is any reflection or self-awareness, just an ever more complex collective of chemical reactions.

The ***fifth***, and most distinctive, aspect of human subjective experience is self-awareness itself: our ability to say, "I am … / I feel… / I see …". It's one thing to see a color or taste a fruit and have a certain reaction to it. It's another to reflect, "Plants, mountains, and some people are unable to see, but *I* am conscious and am seeing 'red'. Or "This landscape is gorgeous. *I* have never seen anything like it before, and *I* will never see it again." Because language enables us to compare our current situation with other possible ones, self-evaluative statements carry additional associations of fortune, regret, pity, self-worth, embarrassment… and their consequent reactions and associations.



44

Mary, The Color Scientist – Color & Self-awareness[45]

Frank Jackson poses an interesting dilemma[46] that any proposed explanation of subjective experience must address. Let's imagine a scientist, Mary, who has lived all her life in a black and white room with a black-and white television. Although Mary has never experienced color, she is an expert on the mechanical perception of color and has read the explanations of color and subjective experience presented here. Supposedly, she now understands "everything there is to know" about color, but if she were to walk outside and see colors for the first time wouldn't she *learn* "something" new? "Something" which hasn't been explained

Well, no. According to the thought experiment, Mary can examine her own mental connections and associations so that she knows before leaving the room what her reaction will be. She knows that with "light blue", for instance, she will have a reproduction/incorporation response of a certain intensity, certain physiological and cognitive changes to which she will also have emotional reactions, visual associations[47] of "immensity", "softness"… Let's also suppose that she has already experienced this reaction before with a certain shade of gray or that she can somehow trigger this exact ensemble of emotions, associations, and physiological changes even if she has never seen blue. (This is actually an assumption that Jackson's dilemma demands. Language isn't just shared words; it is shared or, rather, convergent experiences or associations that are linked to those words. Our scientist has supposedly learned "everything there is to know" about color although she has never experienced it. Yet, it is impossible to learn anything about X, much less know everything about X, without some personal experience to relate to X. If Mary cannot somehow relate what she has studied to her own subjective experience, then our scholar hasn't learned anything.)

Now she walks outside. Isn't something still different or unexpected? Not as far as the ensemble of feelings or emotional reactions are concerned. These she has already experienced before. On the physical level, however, this will be the first time Mary has experienced these reactions being caused by "blue" photoreceptors firing. It would be like being tapped on one shoulder and then another: although otherwise indistinguishable, it will be a unique sensation. She already knew what it "felt like" to see blue, although she had never had the experience of seeing blue. Additionally, on the level of self-awareness, the experience would be unique because for the first time she could say, "*I* am seeing blue" a statement which will itself contain additional associations and emotions of good fortune or regret at not having seen it before, etc.…[48] And this is just a simplified scenario. Of course, Mary wouldn't see blue alone, but would, in fact, be bombarded by a kaleidoscope of new colors and emotion/association ensembles, quite probably (unless she had knowledge of the exact scene awaiting her outside) causing additional emotions of surprise and curiosity.

Free Will. We have no direct knowledge of other minds, so the only philosophical argument for Free Will is one from introspection. Yet our capacity for introspection is constrained by the very things that make it possible: our capabilities of detection, reaction, and linguistic association. This is not to say that introspection is inaccurate or an illusion just that there is only so much that introspection can tell us. The very act of introspection in order to respond affirmatively to the question "Do I have Free Will?" is already dependent upon the assumption that "I" have the necessary knowledge, perspective, and "free will" to correctly answer the question. So, the argument from introspection, the only argument for Free Will actually on the table, is hopelessly circular. We are forced to ignore the constituent parts of countries, corporations, and other organizations when the paradigm of discussion makes no provision for them: "The *United States* or *Company X* decided to…" While such statements may be true as far as they go, in these instances, we also have other paradigms readily available to describe the same events in terms of branches of government and business units, or even at lower levels in terms of citizens and employees. Yet, our capacity for alternative levels of introspection is

limited. That said, the self is not an illusion any more than any other category like 'France' or 'tree'. No frame or category exists in the world independent of the mind; all classifications are dependent on our unique sensations, emotions, and associations.[49] The shadows on the wall are no less real than the objects outside Plato's cave; there is just another way of understanding them. This paper presents an alternative paradigm for the self where our decisions and feelings, like those of other animals, are limited and determined by the prison of our individual instincts and experiences.[50] Nevertheless, through self-awareness and reflection, we can seek new experiences, perspectives, and possibilities and "consciously" force back the prison walls.[51]

Conclusion. Self-awareness isn't an all-or-nothing quality that humans are born with and animals are not, but consciousness--the information that an entity can process about itself and the world--is a continuum with bacteria (or chemical reactions like rust) at one end and human self-awareness at the other. Nor is our continuum the only one imaginable**.** Self-awareness is possible whenever basic information processing systems are organized in such a way as to collectively detect, react to, and associate stimuli as a monad or single unit within an environment of interaction and communication with other comparatively structured individuals. Furthermore, each of the innate abilities we outlined (detection, reaction, and association) is conceivably artificially reproducible. There is no theoretical reason why we couldn't construct an android which, like a baby, was capable of developing self-awareness through experience. Our organs and autonomic responses are unique to species that share our evolutionary history, so our android wouldn't feel its heart quicken and muscles tighten when it was startled (although it might, in a similar way, detect changes in its internal energy level and readiness.) It might not express its emotions through laughter or crying. Nonetheless, it could come to be aware of itself and of its place in the universe in a sense that would be different than--yet still comparable to--our own.

---

[1] Robert Plutchik, *The Emotions* (1991) (Lanham, Mass.: University Press of America)
[2] Antonio R. Damasio (1994), *Descartes' Error. Emotion, Reason, and the Human Brain* (New York: Avon Books,) 149.
[3] Ortony, A., & Turner, T. J. (1990). "What's basic about basic emotions?" *Psychological Review*, 97, pg. 316.
[4] Damasio 102.
[5] Alan Watts (1989), *The Way of Zen* (New York: Vintage Books) 73-4.
[6] Boris Cyrulnik (1995), *La Naissance Du Sens* (Paris: Hachette Littératures) 96-7.
[7] Lesley J. Rogers & Gisela Kaplan (2000), *Songs, Roars, and Rituals: Communication in Birds, Mammals, and Other Animals* (Cambridge, Mass: Harvard University Press) 44.
[8] Rogers & Kaplan 56-7.
[9] Rogers & Kaplan 144.
[10] M. E. Tson (2000), *From Dust to Descartes* 70
[11] Mario Pei (1965), *The Story of Language* (Philadelphia: J.B. Lippincott Company), 105-8.
[12] Damasio 92. The different sensory systems of the brain cannot talk to each other directly but are connected by the higher order association cortices.
[13] Bernard J. Baars & Katherine McGovern (1994), "Consciousness," *Encyclopedia of Human Behavior* (San Diego: Academic Press, Inc.) Volume 1, 691Baars & McGovern 696.
[14] Tson, *From Dust to Descartes* 82. But compare Ludwig Wittgenstein's ideas of meaning as use.
[15] Paul M. Churchland (1988), *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind* (Cambridge, MA: MIT Press) 51.
[16] Cyrulnik (1995) 96-7.
[17] Steven Pinker (1997), *How the Mind Works*. (New York: W. W. Norton & Company) 422.
[18] Rogers & Kaplan 144.
[19] Baars & McGovern 694-8.
[20] Watts 47, 122.
[21] Mary Warnok (1970), *Existentialism* (Oxford: Oxford University Press) 55.
[22] Damasio 240.
[23] Damasio 102-4, 105.
[24] Baars & McGovern 698.
[25] Plutchik 71, 112-115.

[26] Plutchik 71, 112-115.
[27] Plutchik 124-5. There is some disagreement as to the best method of categorizing the basic emotions (See Ortony, supra), but eight primary emotions, each with just four distinguishable levels of intensity and combined in groups of two and three, could produce more than 224 different emotional states. If we assume more levels of intensity this number increases into the thousands - all reducible to eight primary states. No to mention, as we shall see next, the innumerable associations which, in any particular experience, also give rise to emotional reactions)
[28] Damasio 149.
[29] Plutchik124-5.
[30] Ortony, A., & Turner, T. J. (1990). "What's basic about basic emotions?" *Psychological Review*, 97, pg. 316.
[31] Judd, Deane B.; Wyszecki, Günter (1975). *Color in Business, Science and Industry*. Wiley Series in Pure and Applied Optics (third edition ed.). New York: Wiley-Interscience. p. 388.
[32] Baars & McGovern 698.
[33] C.L. Hardin (1988), *Color for Philosophers* (Indianapolis: Hackett Publishing Company) 129; "Colour." *The New Encyclopædia Britannica* (1995) Volume 16: 603.
[34] Hardin, 31-5
[35] Hardin 164.
[36] "Science Does it with Feeling," *Economist* 20 July 1996: 72.
[37] Damasio 147, 162-4.
[38] Damasio 147, 162-4.
[39] Plutchik 15, 48-9, 82, 88-89.
[40] Plutchik 15, 48-9, 82, 88-89.
[41] Charles Darwin (1872), *The Expressions of the Emotions in Man and Animals.* (Chicago: University of Chicago Press, 1965) Chapter III, pg. 77 Chapter XII, pg. 289-291 cited in Plutchik 72-73.
[42] Compare Daniel C. Dennet (1996), Kinds of Minds. (Basic Books, New York) pp. 27-35.
[43] Tson, *From Dust to Descartes* 87-88. Compare ideas from Ludwig Wittgenstein: "One thinks that one is tracing the outline of the thing's nature ... one is merely tracing round the frame through which we look at it." (Tractatus Logico-Philosophicus, 4.5); Immanuel Kant: Our minds and senses do not discover order in the world, they impose order on the world. (See generally Critique of Pure Reason, #27 of Transcendental Analytic); See also Nagarjuna and Zen Buddhist critique of language and subject-object dualism: "objects are also events …our world is a collection of processes rather than entities" "logic and meaning, with its inherent duality, is a property of thought and language but not of the actual world. The nonverbal, concrete world contains no classes and no symbols which signify or mean anything other than themselves…" Duality arises only when we divide and place aspects of our experiences into separate mental boxes. (Watts 5, 73)
[44] See Note 1 for Plutchik's Wheel of Emotions.
[45] Response to Jackson's thought experiment excerpted and abridged from Tson (2000) pg. 97.
[46] Frank Jackson, "What Mary Didn't Know", *Journal of Philosophy*, LXXXII, 5 (May 1986): 291-295.
[47] Without a cultural experience of color, any associations would be limited.
[48] Assuming these weren't included in her pre-color analysis/experience of blue.
[49] See note 43 supra.
[50] Cyrulnik (1995).
[51] Tson, *From Dust to Descartes* 116.