Bernard J. Baars & Katharine A. McGovern

“Self” as an unconscious executive system that receives conscious input and controls voluntary actions.

"... the total self (is) partly known and partly knower, partly object and partly subject ... we may call one the Me and the other the I ... I shall therefore treat the self as known or the me, and ... the self as knower, or the I ...

--- W. James, Psychology: Briefer Course. (1892) (p. 189).

Readings: Chapter 7, Theater.

See also:
http://www.personalityresearch.org/bigfive.html
Phineas Gage was the first reported patient with major brain damage that seemed to selectively impair his personality, his “executive self.”

The following comes from a website written by a historian:
http://www.deakin.edu.au/hbs/GAGEPAGE/

“Phineas Gage was the foreman of a railway construction gang working for the contractors preparing the bed for the Rutland and Burlington Rail Road near Cavendish, Vermont. On 13th. September 1848, an accidental explosion of a charge he had set blew his tamping iron through his head.

(Note: A tamping iron was used to tamp down a charge of dynamite in a blasting hole. At the time, dynamite was very unstable and could easily blow up from a spark coming from the tamping iron striking another metal object in the hole).

The tamping iron was 3 feet 7 inches long and weighed 13 1/2 pounds. It was 1 1/4 inches in diameter at one end (not circumference as in the newspaper report) and tapered over a distance of about 1-foot to a diameter of 1/4 inch at the other. The tamping iron went in point first under his left cheek bone and completely out through the top of his head, landing about 25 to 30 yards behind him. Phineas was knocked over but may not have lost consciousness even though most of the front part of the left side of his brain was destroyed. Dr. John Martyn Harlow, the young physician of Cavendish, treated him with such success that he returned home to Lebanon, New Hampshire 10 weeks later.

Some months after the accident, probably in about the middle of 1849, Phineas felt strong enough to resume work. But because his personality had changed so much, the contractors who had employed him would not give him his place again. Before the accident he had been their most capable and efficient foreman, one with a well-balanced mind, and who was looked on as a shrewd smart business man. He was now fitful, irreverent, and grossly profane, showing little deference for his fellows. He was also impatient and obstinate, yet capricious and vacillating, unable to settle on any of the plans he devised for future action. His friends said he was ‘No longer Gage.’”
The tragedy of Phineas Gage would only be a story were it not for two facts. First, some essential evidence has been preserved: His death mask (an impression of his face after death, probably using gypsum), his skull (now at a Harvard museum), and the tamping iron that was driven like a bullet through the front of his brain. Using computer reconstruction techniques, Hanna Damasio and co-authors were able to reconstruct the likely location of the brain damage (Damasio et al, 1994). (See the figure below).

Second, many hundreds of patients with similar damage have now been studied, and it is believe that the prefrontal cortex, and perhaps its connections with the right temporal lobe, are involved in the brain basis of “self” functions, like personality. (You can search PubMed under “prefrontal brain damage” to get a sampling.)

For us, the key point is Phineas Gage’s change in “who he was” in the eyes of other people. Beginning as a conscientious, goal-directed, hard-working person before his accident, he became “… fitful, irreverent, and grossly profane, showing little deference for his fellows. He was also impatient and obstinate, yet capricious and vacillating, unable to settle on any of the plans he devised for future action. His friends said he was ‘No longer Gage.’”

Conscientiousness, the ability to be persistent in the face of obstacles, is one of the five major personality factors that has been found many times, over more than 50 years of personality testing. It is one of the so-called Big Five Personality Factors. As we will see below, the related experience of mental effort seems to involve quite specific parts of the prefrontal cortex.

Thus we have evidence from the life of Gage, including physical evidence, and even more convincing, from many similar patients who have been studied carefully. In the last ten or fifteen years we have also been able to look at “executive brain functions” in the living brain using techniques like fMRI. This is unique historically --- brain damage throughout history could primarily be studied post-mortem. So we are living in a remarkable historical moment for beginning to understand the self and its basis in the brain.

Below is Damasio et al’s illustration of the Phineas Gage’s brain damage. (Figure 9.1) Notice that the damage, though extensive, was localized --- as far as we know. However, just as when we are injured in any other part of the body, brain damage creates a lot of bruising, swelling, inflammation and cell death outside of the directly affected areas. In Gage this may have been mitigated by the heat and friction of the tamping iron, which may have cauterized the neighboring areas, minimizing bleeding and infection, and sparing neighboring brain cells. In 1848 there were no antibiotics, and most people with injuries like this would have died. The main reason why we think the damage must have been narrowly confined is Gage’s rapid and apparently rather complete recovery – with the exception of his change of self. Gage was a lucky man, in some ways.
Now take another look at the William James quotation at the top of this lecture. "... the total self (is) partly known and partly knower, partly object and partly subject ... we may call one the Me and the other the I ... I shall therefore treat the self as known or the me, and ... the self as knower, or the I ... " (p. 189).

To see this point, ask yourself, when his old friends thought that Phineas Gage was a different man, did he experience that himself? We don’t know the answer, of course, but in many cases like this seem to have no idea that they have changed. They tend to explain their changed behavior in terms of changed circumstances. That is one piece of evidence suggesting that we typically do not have conscious access to our own personality or executive functions. IN William James’ words, to ourselves we are not an object of experience --- unlike that computer screen in front of you, which is an object of experience. Rather, we are a subject, or as we would prefer to say, an unconscious contextual system. (See the Lecture 8 and Chapter 6). Another way of saying that is to say that the “self” is a perspective, or a point of view, which does not enter consciousness as a specific experience we can point to, but which “holds” other conscious experiences and shapes them.

As James says, we are both “I” (the subject) and “Me” (the object) of experience. We can obviously think about ourselves, and have a conception of who we are. However, human beings are notoriously inaccurate in describing ourselves. As the Scottish poet Robert Burns wrote,

Wad that some power the giftie gie us
Ta see oursels as ithers see us

It’s difficult or impossible to do so with the kind of accuracy we have in describing a visual scene. Psychologists have noticed the same thing. There are many conscientious people who constantly have the feeling that they are not working hard enough to meet that next deadline. They may blame themselves for not being conscientious enough --- which may be a good indication of high conscientiousness.

That does not mean that conscientious people can never understand that they are harder-working than others. Rather, at the very moment when they are most conscientious chances are that they are not aware of it as a governing force in their lives. One common sign of conscientiousness is the tendency after succeeding at one major goal to choose another, even more challenging one. If you have only one Nobel Prize, what do you do next? Or one Olympic Bronze medal? At such moments people rarely tell themselves that they are too conscientious – rather, they just focus on finding another hurdle to jump. That is what we mean by saying that personality factors are contextual --- they are unconscious at the moment when they shape our conscious experiences and goals.
Notice, by the way, that we are not saying that the self is unconscious because it is repressed. There may well be aspects of oneself that are repressed or suppressed, but there is no reason to think that we suppress the fact that we are conscientious. From our point of view, there is a great deal that is simply contextual about ourselves --- unconscious because it has become so predictable that we simply take it for granted. It is the active background of conscious experience.

Here are two sets of brain images, one involving “executive regions of cortex” and the other, “consciousness-related regions of cortex.” Take the first, involving “executive control.” (Figure 2, below)

Figure 2. Cortical regions activated in a number of different studies of executive control. From Frackowiak et al (2004) Human Brain Functions. Elsevier.

Left Hemisphere seen from left. Right hemisphere seen from left.

Notice that prefrontal cortex was found to be highly activated, on both the left lateral surface, and the right medial surface --- the inside of the half grapefruit. Since these figures represent many different studies, they seem to confirm the story of Phineas Gage --- that executive processes --- like the brain regions that make us work hard --- are concentrated in prefrontal regions.

Now take a look at Figure 3, below. This figure represents six different experiments in which conscious and unconscious perceptual events were compared --- like the masking studies we discussed in Lectures 2 and 3. That is, the figure shows the parts of the brain that become
active when sensory input is conscious --- even though the main sensory regions are not differentially activated. So they seem to be the parts of the brain that are needed for conscious experience.

The figure shows two dashed circles that capture most of the “hot spots” in these studies. One is in the rear half of the brain, in the parietal cortex. The second is in the front half of the brain, probably including both frontal and prefrontal regions. (Prefrontal cortex is actually closer to the eyes than frontal cortex, a confusing terminology). What is really interesting is that Figure 2 and Figure 3 have a lot of overlap in the front half of the brain --- the part associated with executive functions.

Our basic claim is that “The self is an unconscious executive system that receives conscious input and exerts control over voluntary actions.” That’s the title of this chapter. This kind of brain imaging evidence provides evidence for that idea. The frontal/prefrontal regions are apparently needed both for normal sensory consciousness and for executive control.

What about the other dashed circle, the one in the rear half of the brain, in the parietal cortex? In the last lecture we discussed evidence that parietal cortex involves egocentric contextual maps of the sensory surround --- that is, the point of view of the observer of conscious inputs. (It has other functions as well, but we’ll stick with this one). One reasonable idea is that the parietal hot spots involve that aspect of self --- the self as an observer, as William James called it.

It seems that we cannot understand consciousness without some reasonable conception of self (e.g. Tulving, 1985; Dennett, 1978). "Self" is much like consciousness, a core psychological question that stubbornly survives all efforts to circumvent it. Self-other differentiation is a central concern in sensorimotor control, in mother-child interaction, in the
development of personal autonomy, and even in the immune system (Edelman, 1989).

There is of course a major literature on the psychology of self and other, with experimental, clinical and philosophical contributions (e.g., Freud, 192.; Hartmann, 1958; A. Freud, 1938; Kihlstrom & Cantor, 1984). We will sketch some ways GWT makes contact with this literature.

Daniel Dennett has written "That of which I am conscious is that to which I have access, or (to put the emphasis where it belongs), that to which I have access" (1978). This certainly fits our everyday language. All English sentences about conscious experience use personal pronouns, as in "I see a book," "You are imagining that pain," etc. Certainly we would be surprised if we could not become conscious of some vivid recent memory, some sight, smell or taste in the environment, or some well-known fact about our own lives. The "self" involved in conscious access is sometimes referred to as the self as observer. Similarly, commonsense statements of voluntary control use "self" as an agent, as in "I told him to go," "He decided to find out more," and "I am responsible for my own actions." Again, we would be upset if we were unable to move an arm, stop an intended speech act, or control a usually manageable conscious craving or impulse. The controlling agency for this expected domain of voluntary control is sometimes labeled the self as agent (e.g. James, 1890). Skepticism about controlling agency and observer – but we can now see them in brain scans.

Ryle’s Red Herring.

Some philosophical readers may be surprised by the hypotheses explored in this chapter. It has been an article of faith in analytic philosophy that conceptions of self are vitiated by the “homunculus paradox,” Gilbert Ryle’s idea that an observing self must necessarily contain another observing self, and so on ad infinitum. The self therefore seemed to lead to an infinite regress, an absurdity. But Daniel Dennett, a famous student of Ryle, has provided a straightforward answer: There is nothing wrong with postulating an executive self, as long as it is not composed of little selves. That’s the move that gets us into an infinite regress. In real life we see all kinds of executive systems that guide actions, in Mars Rover robots, for example.

Using the theater metaphor, we can think of self as the director behind the scenes. In the jargon of global workspace theory, it would be called the “dominant goal context hierarchy.” But “stage director” is fine for our purposes.

Losing a sense of self.
Remember that there is a useful distinction to make between the self as a largely unconscious executive (the “I”), and the parts of ourselves that we can become conscious of (the “Me). But we don’t mean to dismiss our “sense of self” --- which is fundamental in emotions like pride, shame, self-confidence and so on. This becomes especially important, because people sometimes complain about “self-alien” experiences, including a loss of a sense of self.

In everyday life, self-alien events occur with slips of speech and action that are disowned (“I didn’t mean to say that!”), with unexpected events like auto accidents that lead to brief periods of shock and mild dissociation, experiences of betrayal of valued social obligations, and in many similar cases of fundamental mismatch. Young parents who wake up at night with a crying baby often report feelings of anger and rage that they perceive as very inappropriate. Most do not act upon such feelings, and may be very upset at themselves for even having them. That sense of upset --- that “I am not the kind of person who feels rage at a crying baby” --- is a sense of mismatch with one’s expectations about oneself. Reliable evidence is available about these situations, and they are all grist for our mill. We develop a contrastive analysis based on it below.

More seriously, self-alien experiences are often reported in disorders like depersonalization, psychogenic fugue, and multiple personality (9.xx). Indeed, all of the Axis-I disorders of the psychiatric manual are marked by self-alien symptoms including unacceptable feelings, actions, thoughts (see DSM-IV), as are many neurological disorders like alien-hand syndrome.

*Avoiding self-knowledge to protect self-esteem.*

Perhaps the most obvious reason to differentiate between the self and self-concept are the extent to which we profit from self-deception (Holmes, 1972, 1974; Goleman, 1985). Even critics of the concept of repression believe that inconvenient thoughts are often willfully avoided and suppressed (Holmes, 1978). Wishful thinking (and sometimes catastrophic thinking) stands in the way of an accurate self-concept. In everyday life we often surprise ourselves with unexpected feelings, actions, inner speech and mental images. If we knew ourselves completely this might not happen. Accurate self-knowledge seems to be unusual, unless of course we are forced to confront something like a broken leg; self-knowledge is always incomplete.

*How to lose touch with oneself.*

A large number of studies show that people often make false claims about themselves when there seems to be little motivational pay-off in doing so. These studies come from two streams of investigation, one focused on errors of attribution, and the second on cognitive dissonance. In a typical attribution study a subject may be led to believe that the sound of an accelerating heart-beat is the sound of his/her own heart. This false feedback has been shown to affect subjects’ perceptions of their own emotional excitement, for instance in the context of
threatening stimuli (e.g. Valins & Ray, 1967; see also Schachter & Singer, 1962). A variety of cognitive dissonance studies show that people judge an unpleasant event more as more positive if they are given inadequate justification for their involvement in it. Presumably they create a higher evaluation justify their involvement after the fact (e.g. Zimbardo, Cohen, Wisenberg, and Firestone, 1969). In general, these studies show that people are pretty inaccurate in their descriptions of their own motives.

Obviously that isn’t always true. When we haven’t eaten for a day we feel genuinely hungry, and we really do like to feel loved, or things like status, prestige, wealth, fame, fortune, or whatever Brownie Points we pursue in life. We really don’t like pain and we really do like pleasure.

So what’s different about the social psychology studies we just mentioned? Nisbett & Wilson (1977) claimed on the basis of such studies that human beings have no privileged access at all to their own processes. This conclusion has been criticized as overstated (e.g. White, 1982; Ericsson & Simon, 1984). The critics have pointed out that there are many examples of accurate introspection, and that we really need to know under what conditions we can expect accurate judgments about ourselves.

Nevertheless, the evidence remains strong that people cannot tell us about themselves much of the time, even when common sense would expect them to be able to do so. For example, when people choose from a display of identical stockings, they will tend to choose the right©most, or the best™illuminated stockings. Asked to explain their preference, they will produce plausible explanations with an air of conviction; but they will not know the reasons for their action. In general, human beings greatly overestimate the accuracy of their self-knowledge, and seem to freely fabricate answers about their intentions, especially when the information available to make the judgment is inadequate.

*Our self-concept is often very oversimplified.*

Another reason for doubting the identity of self-concept and self is the extraordinary oversimplification that seems to characterize our self-concept. The self-concept seem to be value-laden, reducing the complexities of life to very simple "shoulds" and "wants." These beliefs about ourselves often seem to be context-free and absolute. "I'm a likable person. I'm hard-working and want the best for everybody. I have a lot of friends." In contrast, the organization of the self-system seems to be complex, multilayered, sensitive to circumstances and adaptive. In the overall self-system, the self-concept might play a monitoring or supervisory role.

Research on thought monitoring by Singer and coworkers suggests one explanation for the surprising oversimplification of our self-concept (e.g. Pope and Singer, 1978). If we were to
track every bit of inner speech produced by one person, day after day, we would quickly fill volumes. The stream of consciousness is lengthy, jumpy, often self-contradictory, and complex. When we are asked to describe ourselves, we are forced to summarize this vast lode of information. People are generally unable to produce accurate summaries for great amounts of diverse information (e.g. Tversky & Kahneman, 1973; Newell & Simon, 1974; Ericsson & Simon, 1980).

This does not mean that it is hopeless to ask people about themselves and expect accurate answers. Rather, it is vital to do so under optimal conditions, and not to expect people to have access to the deeper layers of their self-organization.

The boundaries between self and the world.

When the eyeball is gently pressed with a finger, the world seems to jump; but, as Helmholtz noted in the 1860's, this does not seem to happen with normal eye movements (Helmholtz, 18xx/1962). Evidently the visual system can distinguish between self-generated and externally induced movements. Somehow self-generated movements are compensated for, so that the experience of the world remains stable in spite of our movement. Self-other differentiation is absolutely necessary not just in the visual system, but in any sensory system, natural or artificial. If a radar dish rotates at a regular rate and detects an apparently moving object, it must differentiate between movements due to its own motion, and those that are due to object itself. Otherwise the moon can be interpreted as a rapidly moving object, and a passing flock of birds as a stationary object in space. Thus self-other differentiation is fundamental indeed, even in perceptual-motor systems.

One can easily show the same need for self-other differentiation in the social world, or in the realm of self-evaluation and personality; to explain all the ways in which mark boundaries between ourselves and the world, we need a concept of self as a multilayered entity with perceptual-motor, social, personality, and other components. We will focus here on the personality realm, contrasting self-attributed and self-alien experiences, but with the clear understanding that these layers of the self system cannot ultimately be separated. Amputation of a limb will impact perceptual-motor processes most directly, but it may create major changes in the socially defined self as well.

Self-alien experiences

Just as we have contrasted comparable conscious and unconscious processes, we can also compare cases where "self" is perceived as "not-self," or as "another self." The most radical,
cases involve psychogenic fugue, multiple personality, and depersonalization disorder. The standard psychiatric diagnostic manual DSM III is an authoritative source on these conditions (Spitzer, 1979). We will briefly review all three syndromes.

If self can reasonably be viewed as a stage director, a very complicated contextual system, we can make some predictions about the effects of violating it. These violative events may be either internal or external in origin, but we should certainly expect "shocking" external events, which are easy to observe, to trigger a disruption in the stage director. Fundamental life changes sometimes evoke surprising alterations in thought, images, inner speech, feelings, and action.

Depersonalization disorder. DSM III describes this as "an alteration in the perception and experience of the self so that the usual sense of one's own reality is temporarily lost or changed. This is manifested by a sensation of self-estrangement or unreality, which may include the feeling that one's extremities have changed in size, or the experience of seeming to perceive oneself from a distance. ... the individual may feel 'mechanical' or as though in a dream. Various types of sensory anestheasias and a feeling of not being in complete control of one's actions, including speech, are often present. All of these feelings are ego-dystonic (self-alien)...." (p.259) Mild depersonalization is quite common: it is estimated to occur at some time in 30% - 70% of young adults.

Depersonalization has many of the expected features. First, it is often triggered by severe stress, such as military combat or an auto accident, physical pain, anxiety, and depression. A similar syndrome can occur after brain-washing, thought reform, and indoctrination while the captive of terrorists and cultists --- all cases in which routine, dominant goals and perspectives are profoundly challenged. These facts are consistent with the notion that disruption of the self involves deep context-violation. Indeed, stress may be defined as a deep violation of expectations and intentions (goal and conceptual contexts) (ref. Horowitz, 1976). Onset of depersonalization is therefore likely to be rapid, as is indeed found, while recovery may be slow, because it takes time to reconstruct a disrupted fundamental context. The high incidence of depersonalization in early adulthood is also significant, since people often establish their fundamental goals and expectations during this period in life, while at the same time going through major life changes which may challenge a new, tentative integration.

Second, there are evidently changes in the way victims of depersonalization experience themselves and the world, consistent with the fact that contexts constrain conscious experiences (4.xx). Along these lines, DSM III states that "derealization is frequently present. This is manifested in a strange alteration in the perception of one's surroundings so that a sense of the reality of the external world is lost. A perceived change in the size or shape of objects in the external world is common. People may be perceived as dead or mechanical... Other associated features include ... a disturbance in the subjective sense of time." (p. 259) Evidently, as the self is challenged, the perceived world may also be estranged.
Psychogenic fugue provides another example of a self-alien syndrome. It involves "sudden, unexpected travel away from home or customary work locale with assumption of a new identity and an inability to recall one's previous identity. Perplexity and disorientation may occur. Following recovery there is no recollection of events that took place during the fugue." This diagnosis is not made if there is evidence for organic disorder.

Again, this disorder can be seen to be a result of deep violations of the normal dominant context, followed by an effort to create a new dominant context, free from the environment that created insupportable problems for the original identity. Fugue typically "follows severe psychosocial stress, such as marital quarrels, personal rejections, military conflict, or natural disaster." (p. 256) It seems to be related to i̞psychogenic amnesia, in which a loss of memory occurs after severe psychological stress. Amnesia sometimes involves a loss of personal identity, but no purposeful travel, and no assumption of a new identity. Note, by the way, that we have encountered spontaneous amnesia before, in our discussion of highly hypnotizable people, who often have spontaneous amnesia for the hypnotic session (7.xx). This is consistent with the notion that high hypnotizables enter a deep, absorbed state, in which they are guided by a context that differs radically from their post-hypnotic context, so that there is relatively little in the way of recall cues available to them afterwards. The issue of spontaneous amnesia and loss of autobiographical memory is indeed a key to the notion of self we are developing here.

The most famous example of self-altering pathology involves multiple personality (now officially labeled “dissociative identity disorder”) (James, 1890; Spiegel, 1984; Hilgard, 1977). Here, too, an eclipsed personality reports a gap afterwards in the flow of experience, just as do victims of amnesia and fugue. "The essential feature," says DSM III, "is the existence within the individual of two or more distinct personalities, each of which is dominant at a particular time. Each personality is a fully integrated and complex unit with unique memories (sic!), behavior patterns, and social relationships that determine the nature of the individual's acts when that person is predominant. ... Studies have demonstrated that different personalities may have different responses to physiological and psychological measurements. One or more subpersonalities may report being of the opposite sex, of a different race or age, or from a different family than the original personality ... The original personality and all of the subpersonalities are aware of lost periods of time..." (italics added, p. 257). Sub-personalities may hear each other or speak to each other, but often with a sense that the voice heard is self-alien --- outside of the self of the current dominant personality (Psychiatric Clinics of N. America, 1984).

Again, we can make an argument for a causal role for some basic challenge to the normally dominant context of intentions and expectations. Thus it is said that "transition from one personality to another is sudden and often associated with psychosocial stress." (italics added, p. 257). Spiegel (1984) has made the case that multiple personality syndrome is

 invariably associated with a history of severe traumatic abuse in childhood. He suggests that children learn when abused to enter a radically dissociated state, which develops over time into an complete, differentiated self. It is easy to interpret these ideas within our current framework.

Recent work with multiple personalities indicates that there is often a "regulator personality," one that keeps track of and mediates between other sub-personalities (Psychiatric Clinics of N. America, 1984). Spiegel suggests on this basis that the normal self, too, may function as a regulator, integrating experience across different situations (personal communication, 1986). This again is consistent with the notion of the self as a dominant context, one that creates continuity across subordinate contexts.

Note the repeated theme of gaps in autobiographical memory in the self-alien syndromes. Autobiographical memory is of course self-attributed experience, and in a GW framework, if self is identified with deep context, we know that it must shape and select characteristic experiences.

It seems as if the more we rely upon something in dealing with the world --- a person, a community, an assumption, a personal ability, or a goal --- the more its impairment will lead to self-alien experiences. This may be especially true when what we have lost is contextual, so that we no longer even know what it was. If we have assumed all of our lives that we can trust people completely, a violation of trust will have ripple effects that spread widely throughout our selves and our experiences. If we rely less on trust, violations of trust might not be so disruptive.