KNOWING OURSELVES BY TELLING STORIES TO OURSELVES

by John A. Teske

Abstract. Part of the epistemological crisis of the twentieth century was caused by empirically establishing that introspection provides little reliable self-knowledge. While we all have full actual selves to which our self-representations do not do full justice, we focus on the formation and existence of a narrative self, and on problematic reliability. We will explore the cognitive neuroscience behind its limitations, including pathological forms of confabulation, the generation of plausible but insufficiently grounded accounts of our actions, and the normal patterns of narrative creation and checking. The evolutionary logic of self-deception may produce adaptive results, particularly in service of the “commitment strategies” that give our species results otherwise unobtainable. It is largely in our close relationships with other human beings, the relationships so well served by these very strategies, that we may find the powerful counterbalancing feedback which may provide positive change and self-transcendence. Nevertheless, we will also warn about a shadow side for which religion can provide both acknowledgment and hope.

Keywords: cognitive neuroscience; confabulation; embodied cognition; epistemology; introspection; narrative; self-deception; self-knowledge; self-representation

John A. Teske is Academic Fellow of the Institute on Religion in an Age of Science (IRAS) and a Fellow of the International Society for Science and Religion; e-mail: john.alfred.teske@gmail.com.
The life of a person is not what happened but what he remembers and how he remembers it.

—Gabriel Garcia Marquez

GNOTHI SAEUTON: LIMITS TO HOW WE KNOW OURSELVES

The Epistemic Revolution

The key to William Shakespeare’s genius, at the turn of the seventeenth century and the dawning of the scientific revolution, was in exploring the drama of characters lacking self-understanding, being unaware of their own motives, just as he explored the hidden processes of interiority in his soliloquies (Greenblatt 2004). Hamlet’s “divinity that shapes our ends, rough hew them how we will” (Hamlet 5.2.11–12), is no less modern than Norman Mailer’s: “Consciousness, that blunt tool, only bucks in the general direction of truth; instinct plucks the feather.” No doubt but that one of Sigmund Freud’s contributions to our own most recent century was to foment an epistemological crisis in his demonstration that Cartesian introspection is neither privileged nor incorrigible (Flanagan 1991). We really do not and perhaps cannot know ourselves from the inside. If our ego defenses form the boundaries of who we are, and they are in fact just a catalogue of self-deceptive tactics for hiding, shaping, or disowning our experience, then the very formation of the ego is based on a complex set of distortions of the truth. Randolph Nesse (Nesse and Lloyd 1992; Nesse 1999) has argued that such defenses are adaptive tactics that are part of our evolved psychological mechanisms. If our conscious self-representations are constructed from these, then our conscious account of our own experience is not likely to be the best guide to who we are, and a hermeneutic of suspicion must be applied to our self-understanding.

The Empirical Failure of Introspection

A lengthy research tradition in psychology began with Richard Nisbett and Timothy Wilson’s (1977) review entitled “Telling More than We Can Know: Verbal Reports on Mental Processes.” They review dozens of studies on cognitive dissonance, attribution, the inability to report the weighting of factors in complex judgment, classic work on the absence of awareness of stimuli demonstrably affecting problem solving, and the reporting of demonstrably ineffective stimulus factors. They conclude that we may have no direct access to higher order mental processes of evaluation, judgment, problem solving, or the initiation of behavior. Most of what is present to consciousness are the results of such processes. Asked about how we perceive or remember, we are clearly stumped, though we have all sorts of beliefs about why we behave and why we like or dislike particular objects or
persons. For example, after participating in experiments in which factors of which they were unconscious demonstrably affect their behavior, subjects deny any such influence. They would respond that such factors might have affected someone else, but not them. Such factors included simultaneous associations (subjects for whom “ocean” and “moon” were paired were more likely to report “Tide” as the name of a detergent), right-left position effects (the rightmost object in a random ordering of consumer products is chosen by about a 4:1 ratio), and anchoring effects (where the use of a “randomly chosen subject” strongly influences the evaluation of average behavior). Subjects denied influences in each case. Nisbett and Wilson suggest that while people do sometimes tell more than they can know, subjects have no “privileged access,” as reports of subjects match those of naïve observers. The source of the reports appears to be the presence of plausible a priori causal theories. This may involve following cultural rules, folk psychologic implications, theories of empirical covariation, and connotative relations. These judgments are not always wrong, but they do not come from or provide introspective access. We also may have knowledge of autobiographical facts superior to an observer, sometimes be aware of intermediate products, have knowledge of prior idiosyncrasies, or have knowledge of attentions and intentions. These do not entail introspective access to mental processes. “It is frightening to believe that one has no more certain knowledge of the workings of one’s own mind than would an outsider with intimate knowledge of one’s history and of the stimuli present at the time the cognitive processes occurred” (Nisbett and Wilson 1977).

What then becomes of the “introspection” that seems to be valued so much by many religious traditions? Perhaps we need to drop the Cartesian model that introspection is what gets you to the hidden, private self. I think the question of how we construct and internalize a narrative for ourselves is far more interesting. To paraphrase Wittgenstein’s famous line about Christianity, if religion is about anything, it is about changing your life, which is what a new story might do. We are all constructing little scenarios in our heads all the time, and managing the impressions others have of us: so a certain view “comes off.” Except that, as several generations of social psychology have pointed out, we selectively listen to the feedback, listening to what we want to hear and discounting what we do not, so, in some sense, our “acts,” put on to fool others, have as their primary purpose fooling ourselves (Greenwald 1980). The rub, of course, is that fooling yourself means you cannot find out what your true self is. If the true self is just the body and its history, then what the stories we tell ourselves often do is to take us away from, apart from, or in opposition to our bodies, so maybe the “true self” is not this inner, secret sanctum, but just the biography of a body. Hoc est meus corpus. Perhaps what connects the different selves from differently situated narratives is just the shared body. It is our friends and
loved ones from whom we get the most useful checking, which we may learn to interiorize. Perhaps that is a model of how our stories become more coherent over time, for example, from members of a church community asking questions about a story brought from a professional network, or vice versa. I think there is always an “other” in these course corrections. Then there are all sorts of ways our stories can become less coherent as well, by splintering off different part of our lives, different contexts, and different social groups. Even going back to when *Homo sapiens* lost its hair, became naked, we have learned to form self-boundaries by hiding things, from our bodies to our simulations, our play. We can then check ourselves against imagined counterfactuals before acting on them. We get good at hiding things. This is where the self comes from. How do you know the words for your own emotions? You learned as a child, when they were obvious from your face, and only later learned to inhibit your facial expression, to act differently, to hide them, put on your poker face, wear your make-up, put on your armor.

The Adaptive Unconscious

Twenty-five years after Nisbett and Wilson’s seminal work, Timothy Wilson’s (2002) *Strangers to Ourselves* cites a plethora of evidence from social psychology of many of the limitations and outright errors to which introspection is prone. Nevertheless this same research tradition has provided evidence for pervasive and often quite sophisticated mental processes with which we evaluate our situations, set goals, and initiate action, while we are consciously entertaining something else entirely. We can focus our attention on one conversation and block out our surroundings, while still monitoring those surroundings in case anything important happens, which will then come to our attention. We learn some things with little conscious effort, including our native language or how to get around a house in the dark. It is also such adaptive unconscious processes that determine very rapid emotional responses, which may initiate bodily responses or actions prior to any awareness, or give us the intuitions we dismiss that happened to be right all along. Indeed, Wilson cites research evidence that too much introspection can cause confusion or even produce errors in judgment. In one of his studies, couples involved in relationships were asked to either think about their relationships for half an hour or think about something else, and it was the latter group that more accurately predicted the future of their relationships (Wilson and Kraft 1993). It appears that for learning about what you feel or what you are like, it makes sense to pay attention to your actual behavior and other people’s reactions to it. Change may come only from how a new self-narrative may mold your adaptive unconscious if you act in ways consistent with it. You become brave by pretending that you are. There is a role of self-deception in bringing about this kind of
self-fulfilling prophecy. We become brave not just by pretending we are, but also by believing it ourselves. This is one of the kinds of prospective, of prolepsis, of “story grasping” of a different future. I think we rarely do this alone.

Our Full Actual Selves

There is plenty of evidence about the emergence, or embedding, of our beliefs, our will, or our actions in our subconscious minds, or in our neuropsychology, and its developmental shaping across the course of our lives. The point is not only the obvious one that decisions are sometimes made too quickly to be the result of our ponderous and slow conscious thought, but that even the ones that seem to be may still be embedded within subconscious processes, which play a contributing if not determinative role. As long as we are willing to surrender any insistence that our agency always be a punctate, moment-by-moment, conscious controller, we allow some role for the subconscious and preconscious processes in which it needs be embedded. This position also would provide a remarkably rich vein of exploration for an ethical view in which consciousness is not the sole carrier of responsibility, in which our “full actual selves,” to use Owen Flanagan’s (1992) apt phrase, includes much more than that of which we are currently aware, or even of what we can become aware. Indeed, this view is happily consistent with an ethic in which the development of virtuous habits, and their socialization in our peers and dependents, also becomes an important ethical consideration, particularly given the vast array of human action, which must become nigh automatic in the functional hierarchy of complex human actions. It therefore becomes all that much more important to address the role of bodily experience in conscious feeling states and embodiment in decision making (Damasio 1999).

Savoir Faire: Knowing from Doing

Embodied Cognition

There is a fast-growing literature in “embodied cognition,” the study of how cognitive processes are aided by actual bodily activities, or their simulation, reviewed extensively for *Zygon: Journal of Religion and Science* (Teske 2013). Since Warren Brown (2017, this issue) gives more attention to this variety of “knowing”; we need not recount it here, short of noting that much of the thinking we do is actually done as much with our bodies, and the nervous systems coextensive with them, than with what is located solely inside our skulls. Our bodies are controlled by our brains, and there is plenty of thought that may reuse the same circuits in accomplishing cognitive tasks. Such tasks are not learned by disembodied minds, but are
housed in and dependent on bodies, both for learning and maintaining knowledge, whether implicit or explicit. Sian Bielock, a psychologist with doctorates in both psychology and kinesiology, has written a very accessible introduction to this research entitled *How the Body Knows Its Mind* (2015). She addresses the role of gesture in memory retrieval, posture in attitude and alertness, and suggests that brains do not distinguish between body and mind. Our bodies may actually “hack” our minds, our movements affecting our thoughts, decisions, and preferences. We use our bodies as learning tools, and in general the body and our immediate environments shape our thoughts, feelings, and actions. Pacing can improve creativity, walking in nature can improve concentration skills, Botox can reduce depression, and physically fit children do better in school.

**Extended Cognition**

There is yet another area of research in “extended cognition,” data showing that the extension of the cognitive systems does not even stop at body boundaries (also cf. Teske 2013). Andy Clark’s *Natural-Born Cyborgs* (2003) describes human beings as human-technology symbionts, and suggests that we have been that way since the invention of words. What distinguishes our long developmental dependency and our neuroplasticity is that we are able “to enter into deep and complex relationships with nonbiological constraints, props and aides” (Clark 2003, 5). Pens, paper, wristwatches, notebooks, calculators, cell phones, and Internet access are just the most recent layers of our extended cognitive systems, and include wide temporal and spatial extension. “Offloading” cognitive functions to calculators, written text, and clocks alters our brains over the course of our lives, and has a long history. Why limit the understanding of mind and person to our boundaries of skin, as “our sense of self, place, and potential are all malleable constructs ready to expand, change, or contract at surprisingly short notice” (Clark 2003, 33)? Anyone who has ever awakened wondering about the ownership of a benumbed hand, or felt personally violated by incursions on one’s property, understands this implicitly. The neural events, conscious and unconscious, that occur inside our skin are also embedded within a social and technological matrix that plays no less a role in knowing ourselves.

**STULTITIAE LAUS: CONfabulation and Self-Deception**

**Narrative, Meaning, and the Narrative Self**

Much of self-understanding is in the realm of symbolic language, as we humans can build a conceptual world of symbols. This influences not only our current experience, but enables us to represent things to ourselves and others that are not immediately present to our senses.
Jerome Bruner (1986; 1990) makes an important distinction between paradigmatic and narrative modes of understanding. The paradigmatic mode involves synchronic understanding via logical proof, empirical observations, and causal explanation, putatively more characteristic of science, and only learned late in the course of human development. The narrative mode, historically more ancient, and developmentally earlier, involves diachronic understanding via storied accounts of “the vicissitudes of human intentions,” organized in time. In narrative, explanations are not causal but are done in terms of believable stories about actors—human and otherwise—striving to do things over time, which is more characteristic of novelist or poets. According to Alicia Juarreño (1999), without narrative, personality traits and human actions are incomprehensible. Therefore, it is necessary to explain them using a hermeneutic, narrative model, as in the case of other interpretive understandings. Causality does not exhaust meaning. A narrative account is not an alternative or opposed to physicality or design but an additional requirement for comprehensibility; it is what we mean by “meaning.”

As Erik Erikson (1958) indicated, becoming an adult involves a reconstruction of the past in a way that leads to the present. Although this does not falsify the past, it must use fictional and imaginative power to “make sense” of the facts as we remember them. “Narrative makes sense of a brain’s own behavior, and may underlie the sense of a unitary self” (Roser and Gazzaniga 2004). To the extent that external, objective events do not occur in storied form, narratives are, from a paradigmatic point of view, always selections, fabrications, constructions and to that extent are always fictional. They can also be told in multiple ways, and by different people. We learn how to tell our stories by first being told stories about ourselves by others, which may form the core of a personal narrative. Stories may include actual events, of course, or fail to do so, and there is a facticity that constrains truth-telling in stories. Donald Spence (1982) distinguished between narrative truth and historical truth, where narrative truth is not the truth of logic, science, and empirical demonstration, but more like verisimilitude. Narrative truth is not about providing external descriptions of the world to be judged by their veridicality. Robert Coles, in his work on the moral imagination (1989), highlights the integrative functions of stories in healing what is sick or broken, bringing together what is shattered, helping us cope with stress, and propelling movement toward fulfillment and maturity.

**Brain Fiction**

There is a problem, of course, in distinguishing between the narrative latitude that produces greater verisimilitude and the violation of what
historians, and, by extension, those trying to hold to the historical truth, call “facticity.” There are a number of neurological deficits that result in confabulation, wherein a patient may invent answers to questions, or accounts of events, fully believing themselves to be telling the truth, like a stroke victim detailing his attendance at a recent conference when he has not left the hospital. William Hirstein, a student of V. S. Ramachandran, argues that such cases are important to understanding the structure of the normal human mind (2005). Rather than admitting ignorance, even normal people can make up an answer to a question or an account of events that is not true, and express it with conviction. In a detailed study of a variety of such cases, Hirstein argues that there is a separation, even on the level of brain function, between the capacity to creatively invent a plausible-sounding story, and a normal checking process that allows us to recognize the story as fantasy.

In split-brain cases where the corpus callosum connecting the two hemispheres has been severed, the language-dominant left hemisphere “interpreter” will confabulate an explanation of behavior produced by information delivered to the right hemisphere (Gazzaniga 1992; 1995). But information may not be freely shared between hemispheres even in normal subjects, the language-dominant hemisphere inhibiting right-hemisphere word recognition (a half-second flash of “tar . get” results in a subject being aware only of the left-hemisphere “get”; Gazzaniga 1983). As we have seen in our discussion of Nisbett and Wilson’s (1977) research in introspective failure, normal people regularly confabulate accounts of mental processes of which they are unaware. A confabulation is primarily understandable as an ill-grounded epistemic claim that the confabulator does not know is ill-grounded, and we all have the capacity to vary the strength of our epistemic criteria, say between talking with friends and testifying in a courtroom.

There is a question of whether there is any final “fact checking” that makes some memories more reliable than others. Every time a memory is brought back to consciousness it gets reconsolidated in the context of recall. Our memories are always and repeatedly reformed in time, so always run the risk of being increasingly altered. I think the case in point here is the research on so-called “flashbulb memories” (Brown and Kulik 1977) of where we were and what we were doing during major public events, for example, the airplanes crashing into the World Trade Towers on September 11, 2001. It turns out that we get all sorts of things wrong in these memories (Talarico and Rubin 2003; Krackow, Lynn, and Payne 2006), even getting the emotional experience wrong 58% of the time a year later. The more often we recall and retell an experience, the less accurate it may be, so perhaps “flashbulb” memories are more confident only because of how often they have been retold, like the “fishing stories” that get better every time they are retold, and we all have our repertoire of favorite stories. On
some level we know we were making them better stories each time we tell them. So memory is all about reconstruction. It is not a record. So in what sense is it “the facts” plus subsequent modifications? It is not, which is why eyewitness testimony, which was once the gold standard in court, is so troublesome. It works this way because it is adaptive, information you learned after the original event gets incorporated into your story of an event. Yes, even when we read, we read for the gist, we are not recorders. But do we retain a core of the “things that transpired” and reconstruct the rest? Or is there a darker possibility that even the “things that transpired” can get altered? James Marcia’s research (1980) on Erik Erikson’s “identity crisis” identifies four different identity statuses based on the dimensions of crisis and commitment in college student’s stories. But is that what happened or is it about how we tell the story? What makes something a “story” is that there is some kind of conflict followed by some resolution or denouement after climactic tension. Ben Slugosky and Gerald Ginsburg (1989) interviewed subjects a decade later and found them saying things like “What crisis? This is the kind of person I am, and I was always headed this way.” Their current “story” is that there was no crisis. Ask students to read from a journal or from emails of a few years back about something that seemed like a major crisis at the time, and they will usually respond with something like “oh my God, I thought that was a crisis? It is nothing compared to what I am going through now.” What is even the “gist” of an event might look different in retrospect. Our current story is always a rewriting of our past, just as each generation rewrites its history. So I think even what seem to be our most reliable memories may be more unmoored than we think. Bodily connections might provide better moorings in our memories of our past, as Odysseus is identified by his bodily scars. Or by what other people remember that can be verified by physical facts. Others might thus be involved in the “checking” of even our autobiographical memories.

Confabulation: Creation and checking. We appear to be generating accounts of our behavior almost constantly, the left-hemisphere language generation system Michael Gazzaniga (1998) calls the “interpreter.” Such an account must be able to pass (or elude) any checking procedures for it to be expressed. Our normal state, in which we might say that we do not know or do not remember something, appears not to be due to nothing being generated, but to the active suppression of that, which is incorrect. There is brain activity that normally suppresses irrelevant memories (Schnider 2001). Knowing we do not know something may only come after generating and rejecting a series of answers. Most of the data (Hirstein 2005) suggest that many of the checking processes involve neural circuits that run through the orbitofrontal cortex. These may include a range of checks, including those mediated by right-hemisphere function such as body
representation and audience monitoring, but they are likely to be bifrontal coordinations, including checks against autobiographical memory, factual knowledge, logical and causal consequences, and reward value.

Failure to pass a check may result in the activation of both orbitomedial frontal cortex and autonomic function. There appears to be an emotional component to most cognitive function, and these emotional-autonomic functions may be important to the formulation of intentions to act, or to inhibit an action. Simply knowing one is in error is not enough to stop making a claim based on it, and failures of autonomic function are common in confabulation. Confabulators have lost an emotional component to thought, showing lower skin conductance responses that might be needed for the negative emotional tags needed to inhibit the formation of an intent or an action. Antonio Damasio’s (1994) “somatic marker hypothesis” suggests that such connections are crucial to the avoidance of risky or dangerous choices that may be based on ill-grounded beliefs. The orbitofrontal cortex’s sensitivity to reward and punishment enables it to help improve reliability by translating from reward values to truth and falsity.

Clues about self-deception. Armin Schnider (2001) suggests that it is the same orbitofrontal processes that suppress memories not relevant to an autobiographical recollection, the brain in general using functions that usually focus and maintain attention. If I can avoid attending to the sound of passing cars, without being aware that I am doing so, I can certainly avoid attending to some emotional difficulty, and also not be aware that I am doing it. The mediodorsal nucleus of the thalamus (damage to which
we also see in the confabulations of Korsakoff’s patients) may function to block sustained cortical activity, to switch between different thalamocortical connections, or to suppress unwanted thoughts. The neurophysiological evidence suggests that evolution may have coopted some areas of the brain to suppress others in self-deception (Anderson et al. 2004). Neural inhibition of one area by another can also be replicated in producing inhibition by direct external electrical stimulation, inhibition of the anterior prefrontal cortex decreasing reaction time, physiological arousal, and moral conflict during lying (Karim et al. 2010). Some classic research shows unconscious recognition of one’s own voice (higher galvanic skin response) under both conscious denial and projection (Gur and Sackheim 1979); conscious self-boundaries contract under failure and expand under success, despite the body’s accurate recognition (GSR response). Interestingly enough, since the limbic level emotional evaluation of a stimulus occurs in the early stages of event encoding (beginning about 120 ms after onset), this permits limbic (emotional) input to “shape the content of the encoded experience rather than just react to it” (Halgren and Marinkovic 1995, 1146).

*The meaning of confabulation.* Gazzaniga (1998) connects his concept of a left-hemisphere “interpreter” to lying and self-deception. He asks about the usefulness of a left-hemisphere “spin doctor” when we are such lousy liars, anxious, guilt-ridden, and sweaty. He suggests that this is because the “interpreter” is what keeps our personal story together, and to do that we learn to lie to ourselves. But the autonomic signs of lying may be connected with the very orbitofrontal processes that provide a check to the explanation-producing process of the interpreter, cast doubt on them, and can prevent them from becoming beliefs in normal circumstances. The function of the orbitofrontal cortex is in the application of standards, social, ethical, and religious, such that we can feel revulsion at products, actions, or emotions that do not meet those standards. With damage or inhibition, we do not produce emotions strong enough to stop us from the inappropriate, the inhibition of which has ethical implications, as the incarnation of our commonsense notions of conscience. William Hirstein (2005) speculates that with intelligence and foresight, confabulation and self-deception are needed to keep at bay the painful truths of our mortality, our insignificance in an immense universe, and the potential for tragedy. Ernest Becker (1973) said the denial of death was the central fact of human psychology.

*Self-Deception*

*The evolutionary logic of self-deception.* There is an evolutionary logic to self-deception, detailed extensively by Robert Trivers (2011) and William
von Hippel and Trivers (2011). Getting around in a socially interdependent world, particularly with the longest periods of childhood dependency of any species (Konner 2012), often involves deceiving others. Hiding things, both from others and from ourselves is an important part of this project, and how we distinguish private from public. Parent–offspring relationships are not conflict-free, and many sex differences can be understood in terms of differential parental investment. Cheaters can easily exploit relationships that might have been reciprocal. In such a social species as ours, it may well be, as Nicholas Humphrey (1976) and many others have argued, that deception provided a powerful selection pressure for the evolution of intelligence, via the arms race of deception and its detection. Hence, we are likely to be pretty good at it.

Why self-deception? Because, as those “practiced at the art of deception” well know, it is far easier to deceive others if you deceive yourself first.

We deny the truth to ourselves. We project onto others traits that are in fact true of ourselves—and then attack them! We repress painful memories, create completely false ones, rationalize immoral behavior, act repeatedly to boost positive self-opinion, and show a suite of ego-defense mechanisms. (Trivers 2011, 2)

This means that the primary function of self-deception is not defensive, operating beneath the toxicities of our delusions of autonomy and independence (in part created by those same self-deceptions), but offensive, its success measured by how well we keep things from others, or fool them.

There is an epistemic cost, of course, to the use of this strategy of deceit by self-deception, in all the ways we are rendered unknowing (or wrong) about ourselves and others. Nevertheless, the biological advantages, in terms of survival and reproduction, are obvious, and the psychological benefit is in terms of feeling better and being happier. While acting on the basis of what is in fact untrue can lead to unpleasant consequences, and alienate us from reality, as long as the cost/benefit proportion weighs more heavily against others and in favor of ourselves the strategy is a winner. We give off fewer of the cues of intentional deception, reduce the cognitive load by remaining unaware of part of the truth, and have an easy defense against detection in the denial of intent.

Self-deception is not a contradiction in terms because our “full actual self” (Flanagan 1992) is composed of many parts. We are well aware, for example, that our self-represented self probably cannot fully be held in consciousness all at one time, which is why we may take a few days to make an important decision. We can also become aware of ways in which our behavior is not consistent with our self-representation, hence commonplace experiences of not feeling or behaving in accord with it ("I just wasn’t
myself`). We all know that we can sometimes actively decide to “not think about something” (suppress) while we focus on another task, as I did when I returned and finished leading a research seminar just after learning that my father had suddenly died. Self-deception may include a wide variety of ways in which we preferentially exclude true information about ourselves or reality more generally, in different degrees of unconsciousness, for varying lengths of time, from momentarily having something “slip your mind” to forgetting it entirely. Not having attended to or encoded something in the first place is likely, of course, to be far more efficient, effectively hiding something from oneself to better deceive others. Even when deception cannot be detected against background behavior, there are often a range of behavioral cues (DePaolo et al. 2003; Vrij 2008). These include nervousness, exertion of control belied by overacting, overcontrol, rehearsed responses, and displacements (Troisi 2002). The most critical cues are the ones due to cognitive overload (Vrij 2004; Vrij et al. 2006).

Trivers (2011) catalogues different categories of self-deception. (1) Self-inflation is the most common, increasing our “benefectance” (Greenwald 1980), what our “totalitarian egos” do to appear more beneficial and more effective. (2) Derogating others is the other side of the self-inflation coin, deflecting attention and producing better relative comparisons, a tendency that has been experimentally demonstrated in circumstances where one’s own capacities have been threatened (Fein and Spencer 1997). (3) In-group/outgroup biases are another part of self-deception (Maass et al. 1989), even in circumstances where the groups are arbitrary and randomly assigned (Tajfel 1982). (4) Moral superiority, the tendency to more harshly judge others than ourselves is common (Batson et al. 1999), although putting someone under a cognitive load appears to leave an unbiased internal judge (Valdesolo and DeSteno 2008). (5) Illusions of control reduce stress and improve performance, though they may actually harm performance when the illusions result in mistakes in judgment, and lack of control gets subjects to see patterns where there are none (Whitson and Galinsky 2008). (6) False internal narratives can also help conceal true motivations and provide a ready-made convincing explanation. In all cases, one of the most common aspects of the self-deceptions that help us deceive others is denial, and our capacity to construct a benefectant public persona to better screen our true intentions and real causes. Part of us may still register accurate assessments of self and other, so there may often be a rather more complex self, separated into public and private aspects that may interact.

Evolution, self-deception, and commitment. Eduardo Gianetti (1997) points out that while self-deception may often be a curse, it is also a source of the commitments we make to futures we cannot know. In this may reside some of the greatest accomplishments of our species, as well as
“the savage, inexplicable hope which feeds us and sustains our lives.” As we shall see below, it may also be important to the personal commitments we make to each other, enabling us to obtain greater goods than could be attained without it, and intimacies that may be the best source of genuine self-knowledge, and of the possibilities of embracing the better angels of our nature.

The evolutionary psychology of moral behavior (Barkow, Cosmides, and Tooby 1992; Wright 1994; Pinker 1997; Buss 1999) provides a powerful story about the deeper logic that may undergird religious faith as well as other human commitments. Nesse (1999) points out that, while kin selection and reciprocity may provide some understanding of moral relationship, even non-zero sum reciprocity (cf. Wright 2000) cannot account for the kind of good provided by deep friendships or life partnerships where help is given when there is nothing to be gained. Nesse calls it a “commitment strategy” involving a kind of “futures trading” that includes commitments to future actions, which would not then be rationally self-interested. Under such circumstances we might realistically hope to get help when we really need it the most, when we are sick, alone, or poor rather than only when we are able to reciprocate. This is the obverse of the same logic of mutual assured destruction that might well have kept the world from nuclear annihilation during the Cold War generation. Why would we believe that our partners would not, in the end, do what was most rationally self-interested, what would give them the maximum cost/benefit gain, and just cut us loose? We believe that they will act in ways beyond self-interest only if the signals of commitment are accompanied by such irrational displays of emotion that we come to believe they would actually follow through. Of course, it might also be in their interest to deceive us about such commitments by themselves being self-deceived, something which our own strategies might well take into account, and they ours. Given that such commitments can provide goods not otherwise attainable, there may be selective advantages to those able to give and receive them, which would provide an evolutionary shaping of the capacity for passionate, emotional commitment. It would also provide the complicated dance of deceptive versions of such expression, of the detection of such deception, and self-deception, which makes our relational lives so poignantly baffling. Nevertheless, our beliefs about the possibility of such commitments are what make them possible; without the ability to give this deep kind of deep trust, one cannot get it. Fortunately, most of us have existent proof of such commitments, if not in the love our parents share for each other, then in what they give us with no expectation of return. No wonder our early attachments are so predictive of early adult intimacies (Hazan and Shaver 1987). People not socialized with experiences of the trustworthy, or who have repeatedly had their trust betrayed, may not be capable of such strategies.
Externalism and Relationality

A number of authors refer to the long history of contributions of the Abrahamic religious traditions, Christianity in particular, to a belief in individuality as separate from others, interior, and bodily restricted (Taylor 1989; Cary 2000; Grenz 2001). Nevertheless, there are alternative strands in these traditions, consistent with an intersubjectively externalist view. These include Karl Barth’s conception of the *imago dei* as existing not in individuals nor in the capacity for relationship, but in relationship itself. Indeed, as Stanley Grenz (2001) has argued, the ascendancy of relational ontologies has been widespread across a variety of theologies since the early twentieth century, and he suggests that it is in our relationships with others, including the bodily and the sexual, extending into our communities, to which our self-understanding might better be bound. Such views are coherent with *externalism* in the cognitive sciences (Clark and Chalmers 1998; McCulloch 2003; Rowlands 2003; Wilson 2004; Noe 2009), the view that “the mind ain’t in the head,” but that heads and bodies are proper parts of minds. Mental phenomena are hybrids of physical events in the head and events in the world to which they are often coupled, not least of which are events both within and between other people and ourselves.

Adult Attachment

Our earliest models of self and other come from the bond of love between caregiver and infant that provides support, protection, and a secure base for exploration, the dynamics of which also play out in adulthood (Bowlby 1969; Mikulincer and Shaver 2007). What we do in situations of danger, upset, and in relationship formation and dissolution tend to reflect secure, anxious/ambivalent, or avoidant attachments, developed initially in infancy, but extended in the ways in which adults engage each other in romantic love (Hazan and Shaver 1987). Adult attachment is less about overt proximity seeking, and more about our working models of significant others (Mikulincer and Shaver 2007). Under threat, our models of attachment are automatically activated to provide the comfort and security made possible by proximity, or by our representation of close others, and memories of helpful comfort in the past. Secure styles are associated with greater relationship interdependence, commitment, trust, and satisfaction. Anxious styles often produce more distress and ambivalence under difficulty. Avoidant styles result in the least amount of helping and support for partners, and report lower levels of intimacy and enjoyment in daily interaction. Securely attached people are far more active and effective in providing secure comfort, encouragement, and support to others.
Our adult attachments may tell us as much about how we see ourselves as how we represent others. This is an implicit knowledge of self and other. Those with secure attachments identify with their attachment objects, and internalize the comforting and soothing qualities of the other. This enables the securely attached to be a secure base and safe haven for others.

*Intimacy and the Boundaries of Self*

Our culture commodifies everything, abstracting things from their context. Steven Winter (2011) points out that we all too often treat the self in the same bounded and commodified way. Winter focuses particular attention on the development of sexual autonomy. Sexual autonomy is not about individual privacy. Winter finds this somewhat surreal: “after all, when one is alone, one does not need a condom” (238). If sexual autonomy is a fundamental aspect of human flourishing, it is because, as Plato points out in the *Symposium*, *eros* is a sexual desire that attaches to a person. It is something that enables us to treat another being as the person they are, sex being an agency by which we respond to each other through our bodies, even at our most carnal, interested in a relationship between persons.

Indeed, in the earliest stages of adulthood, sexuality is the domain in which we learn to be responsive and responsible to the other. The successful negotiation of sexuality and, ultimately, intimacy requires one to develop skills and values such as empathy, negotiation, compromise, cooperation, recognition of and respect for the other. (Winter 2011, 242)

Intimate relationships are one of the primary ways that we seek recognition and establish identity, hence the havoc wreaked upon our sense of ourselves by the pathologies of intimacy, in narcissism, manipulation, and exploitation, or why childhood sexual abuse can destroy the very capacity for agency. An important part of what we expect from intimacy is someone who “sees me as I really am,” and one of the advantages it gives us is the comfort and confidence to be just that, our bodily selves.

*Projection and Reciprocation*

Our relationships are our redemption. We act on each other’s behalf, and show kindness in our bodily presence, with a touch, a kiss of peace, in holding and being held, in assurances of love, in the return of hope, in laughter and in tears. The depth psychologist Robert Johnson (1983, 1991) asserts that “Romantic love is the single greatest energy system in the Western psyche. In our culture it has supplanted religion as the arena in which men and women seek meaning, transcendence, wholeness, and ecstasy” (1983, xi). He suggests that “falling in love” is the most
powerful projection one makes, a projection of our own most noble and valuable aspects onto another human being, and they unto us. The problem is that even if there is divinity in each of us, the projections are not true, and this experience is different from the quieter and more humanly proportioned experience of loving. The intensity of the projection also obliterates the humanity of the beloved, and that while we have loosed the sublime feeling of which we are capable, we “set ourselves up for the greatest suffering we will ever know” (1991, 66). Researchers in psychology regularly distinguish between the “passionate love” of early relationship formation, including frequent thoughts of the other, as well as idealization of the other’s positive qualities and less awareness of their flaws, and the “companionate love” more frequent in longer term commitments, with an emphasis on mutual care, which is related to higher satisfaction in life (Hatfield and Rapson 1993; Diamond 2004; Kim and Hatfield 2004).

What may be most important, in both helping us transcend ourselves, and becoming what our narrative selves would have us pretend, as well as to come to understand ourselves better and more honestly, are our intimate relationships with others, including the enemies that often know us so well. It is in these relationships within which we may be pressed to acknowledge our self-deceptions, by the knowing of others not sharing the same motivations to sustain them, or, by those who do, holding us to the commitments made in the name of those deceptions, and making them real by enacting them. It is in such knowing relationships, where such services are often performed reciprocally, whether for our benefit or theirs, are motivated to press us, gently and lovingly, or dangerously and confrontationally, to become our own better angels, to be more than we were before. It is in these relationships that we both come to know ourselves better, in the “checking” functions of others that we may come to interiorize, and by coming to expect more from ourselves, to do better than we have ever done before.

**Communities and Ethics**

The presence of others is essential, perhaps even definitive of our humanity. Feral, isolated, or severely abused children may never really have the others against whom they can build a sense of who they are, to whom they can be attached, and who reflect back to them. They have no way to construct a richer narrative with purchase on coherence, consistency, or reliability. The construction of a stable narrative incorporates and internalizes, is a product of many others. You can be isolated, living on a desert island, but you still have the accumulation of all the parents, peers, teachers, the people from whom you internalize all of what you consider to be the equipment of yourself. But it is living bodily together with a community that builds a story, and a view of the world, into a person’s life. It is
not just a collection of dyads, but a whole group of people so that the corrections can get corrected, the checks rechecked. Obviously this is a process that can be used in both good and poor ways, as a group can press a story that is pathological, destructive, or violent. The kind of commitment strategies, facilitated by early attachment, that make long-term relationships and marriages possible, also involve commitments to communities, involving some of the same “leaps of faith” or “leaps into pretense” that can have a prolepsis upon a future. One’s marital vows are said not just to each other, but in front of a witnessing community, who will see them vouchsafed. I can pretend to be a particular kind of person, and the result will be that if I do this long enough, get treated this way long enough, that is becomes who I am, and I am no longer fooling myself, and find a particular story important or even central to my work and life. Therapeutic relationships may even encourage such commitments as formative routes to healing.

There are kinds of judgmental communities in which it becomes difficult to grow and develop. We all want communities that are accepting, in which we are appreciated, in which we feel respect and love. But we also need to be held accountable, where respected peers can still hold up a mirror for you to see both your vices and your virtues, not in judgment, but in a way in which you can decide if this is the kind of person you want to be, if maybe you can do better than this. Part of the hermeneutic of suspicion is “don’t be so sure about yourself”; even in spite of the advantages of confidence and the epistemic of accomplishment, there is always this piece that has to say “watch it there,” watch it in other people, but especially in yourself. The other half is the science half. Science is really good at a certain kind of checking: What do you have for data? What is the evidence you have for this? What is the empirical work to support this? But where is the moral checking? Sadly and unfortunately, while lots of scientists are wonderfully moral people, such checking is not so much part of science, suggesting that, while science is necessary, it does not provide a complete world view, as it has little basis on which to decide not how something should be done, but whether.

What is the role that religion, or religiousness, play in the development of self? Part of it may be in not taking a Cartesian view of a self or soul as inside, private, and hidden, or religion as about oneself, but as about an “us,” a chosen people, a community of believers, a society of friends. If we think of ourselves as bodies that act in the world, as part of a church, a synagogue, a mosque, then religion is about a community, expressed outside of me, in the world, and it is not about me. Is not this the ethic of most religions? I do not think it is so much about the salvation of my individual soul. How selfish and egocentric. Maybe it is about giving it away? Perhaps the “more perfect body” of an afterlife is not about the you that you think about yourself as being now, but about this broader model;
is not *religio* about “tying things together”? Is not redemption about healing one’s alienation from a community?

**Nolo Contendere: Knowing What We Do Not Want to Know**

There is a shadow side of ourselves we would rather not acknowledge, but must if self-knowledge is not to devolve to wishful thinking. Failure to address our own darker sides is something about which Enlightenment thinking seems to be in denial. I would like to suggest that this may constitute just one example of the sorts of things that we all too easily miss in knowing ourselves, and do so at our own peril and the peril of others, precisely the sort of thing of which “perilous times” is likely to be constituted, for which our intimate relationships may be the best counterbalance. “Beware the *Dark Side*, Luke.” Ignorance, temptation, and concepts like original sin are important parts of the Western theological tradition, and its devils, but also of other traditions, the tricksters of mythology, of Coyote and Kokopelli, of Pan with his pipes, of Hermes, god of crossroads and thieves. Could not we understand such warnings as about our natural tendencies to think ourselves better than we are?

Terry Eagleton’s (2009) *Reason, Faith, and Revolution* proposes that, at their most convincing, the Jewish and Christian Scriptures have valuable insights into human emancipation, and much to say about vital questions like death, suffering, love, and self-dispossession. There is a common ground between science and religion in the “tragic humanism” that Eagleton draws from theology, Freud, and Marx. The broken body of a political prisoner who has been tortured to death may provide a more realistic picture of human nature than “liberal humanism” in a post-Holocaust, post-Hiroshima era, and a more useful signifier of the human condition. Eagleton holds out for a more nuanced view of religion than one that reduces it to a flawed explanatory system based on unsupported beliefs about a supernatural agent. He suggests that such a view has value in noncircular justifications of rationalism, and may temper the political self-contradictions of an ideology of tolerance and diversity. Most important, from the present perspective, it helps to provide a critique of the self-origination, self-authorship, and self-sufficiency that presume “to pull progress and eventual perfection out of our own entrails” (Fish 2009).

Yes, there is a “dark and troubled side, too.” But the “sunny side” to which religious communities may be built to direct us includes family support, wedding vows, social action, and charity (Stark 1996), just as they direct us away from our illusions about ourselves, and do so with the power of love and forgiveness. It is also clear that the only real power of humanity has been in acting collectively, a particular challenge in an era of fragmented individuality, as argued by Michael Lerner in *The Politics of Meaning* (1997).
and in Spirit Matters (2000). There are also real religious encouragements against self-deception, and our capacity to see flaws in others, like attending to the beam in your own eye before trying to take the speck out of your neighbor’s (Luke 6:42; Matthew 7:3-5). We can also blind ourselves to our virtues, as this means we do not have to take the responsibility for them, but being drawn out by others to take on larger roles, to attempt greater things, to take up our crosses, is no less a part of a loving community. We have a shadow side of which we are often unaware, but we also have a “golden shadow,” with which we get in trouble when its projections onto others keeps us blind to their humanity, and our own overflowing cups, the banquet we may have for others, the surplus of love from which only fear, uncertainty, and lack of faith prevents us from giving freely, and loving wastefully.

**Note**

A version of this article was presented at the 62nd Annual Summer Conference of the Institute for Religion in an Age of Science (IRAS) entitled “How Can We Know? Co-Creating Knowledge in Perilous Times” held on Star Island, New Hampshire, from June 25 to July 2, 2016.

**References**


