Chapter 2
Groupthink and the Evolution of Reason Giving

Check for updates

Gregg Henriques

A short while ago I witnessed what I considered to be an excellent example of "political groupthink" at my university. Because the "truth" of political opinions is inevitably tied to one's own position and political values, I should be clear that my political views are left of center, are consistent with a more classic liberal view than a modern progressive identity political view (see, e.g., Lilla, 2017), and I am concerned that the university academy is too dominated by left leaning political perspectives such that I am a member of the heterodox academy, which is explicitly concerned about the lack of political "view point diversity" in university settings.

The context was a faculty discussion following the disturbing events in Charlottesville, VA, which included a large number of white supremacists marching on the city, carrying Nazi flags, and ultimately committing a horrendous act of violence such that an innocent protester perished. It received national attention, and I had offered a commentary on my *Psychology Today* blog about why I believed President Trump's response that "both sides were to blame" was disheartening (Henriques, 2017). I argued that if we were guided by a clear moral compass, this would not have been the response of our government leader. It was because of the blog that I was asked to be a participant leader in an open faculty discussion about the implications of the incident for college campuses.

What happened, at least from my "center left" political perspective, was that a group of liberal professors engaged in a discussion that proceeded to move from a heavily "left leaning" to ultimately considering an activist stance that, if actually carried forth, I believe would have been an excellent example of the horrible dangers of groupthink. The progression of the discussion was as follows: After a few opening remarks, one professor commented that the conservative commentaries and critiques in the media seemed mean-spirited and off-base, such that the picture being painted about the university academy was "unrecognizable." Then an African

G. Henriques (✉)
Department of Graduate Psychology, James Madison University, Harrisonburg, VA, USA
e-mail: henriqgx@jmu.edu

American professor, one of the two in the group, asked if we wanted to have a "real" conversation about race. She talked about her experiences at James Madison University (JMU), especially since the election of President Trump, and claimed that there were really two cultures at JMU, one for whites another for blacks, and that the JMU culture was racist in many ways. Then a number of historians offered perspectives on the emergence of confederate statues and why they were constructed largely at the time of Jim Crow laws and many were explicitly placed near or around black communities to send a clear message about intimidation and racist attitudes.

Then a professor commented that it is our duty as professors to help students recognize just how racist and sexist our society truly was and is, and how much more needs to be done to level the playing field. Then a self-described "activist" stated that professors were good at talking, but what was really required was action. He argued we needed to march or to demand change. Finally, another activist-professor claimed that what was needed was to send a real statement. He suggested we start a petition to change the name of our university from James Madison to that of a slave who had worked toward freedom. No one objected directly. The meeting was near the end and adjourned shortly thereafter.

One of the main characteristics of groupthink is that, in an effort to demonstrate harmony and unity, people fail to consider alternative perspectives and ultimately engage in deeply problematic decision-making. From my position on the political spectrum, this is exactly what happened in this case. First, there was a general comment that "we" do not recognize the negative ways in which conservatives characterize "us." This created a groupthink element of us versus them and justified the notion that "they" do not know what we are about so we do not need to listen to them. Then claims were made, such as those regarding the institutionally racist culture of the university that almost could not have been challenged, given the context and immediate social dynamics. Then evidence was offered about the racist nature of the culture in general. Then there were proposals about our roles both to educate others about this and to take active steps to dismantle the institutions that emerged historically from our racist past, including cleansing the university of association with the slaveholder, James Madison.

While there certainly are elements of truth and logic to this kind of thinking, the chain of logic was clearly blind to many realities. First, the conversation, its content, process, and evolution of proposals were an exemplar of exactly the kinds of thinking that or about which conservative political theorists object and critique the liberal academy for engaging in. In other words, it was notably ironic that the discussion began with a comment about how a professor claimed the conservative press painted an "unrecognizable" picture of the academy, and then the group proceeded to engage in precisely the kinds of arguments and rhetoric that political conservatives complain about. Second, the progression of stronger and stronger claims, ending with a claim that we should disavow the primary organizing symbol of the university and the father of the United States Constitution. What are the implications if James Madison is deemed to be no longer worthy of admiration?

Third, there were virtually no reflections about the consequences of this proposal. Consider that it was the proposal that we needed to remove the statue of

Robert E. Lee that sparked the initial confrontation in Charlottesville and provided justification for white supremacists and the alt-right for an organized march. Could one imagine the response to changing the name of the university of the Father of the Constitution? In fact, President Trump was roundly criticized for asking, in the context of removing confederate statues, "Who is next, George Washington?" Removing James Madison and replacing it with a former slave would say that we need to completely cleanse ourselves of any association with our "founding fathers." No one in the group, including myself, pointed these things out. All I could muster was the following comment, "I think we would need much more discussion about these issues to determine if there is a consensus in the group about actions that might make sense to take." Why did I not stand up and strenuously object and say that the discussion had clearly gone off the rails? The power of social influence in group contexts is enormously strong.

There were many elements in this situation that made it ripe for a groupthink dynamic. Specifically, we are living in a hyperpolarized political environment and a highly emotional and polarizing event had occurred. The academy is quite liberal in general and this particular event and open discussion would have been appealing to individuals who were particularly liberal. There was a need to "do something" in response and it was important that "we" were united in that response. Indeed, if one was not with the group, then questions would be raised about whether one supported Trump or white supremacy in general.

Social psychologists have long documented the empirical reality of groupthink (Turner & Pratkanis, 1998). However, what is generally not present in such analysis is a deeper understanding regarding the nature of human consciousness and social motivation, such that the reasons why humans are so prone to groupthink are laid clear. Indeed, a general critique I have long made regarding the field of psychology is that it lacks a unified conceptual framework that grounds and ties together the various lines of empirical work that currently defines the field. My scholarly efforts have been devoted to the development of a new, unified theory of psychology (Henriques, 2003, 2011) that pulls together many different threads within the field and related social science perspectives and offers a way to see the whole in a way that is more coherent than the current fragmented arrangement of theories and findings.

The remainder of this chapter focuses briefly on two key ideas that are part of the unified theory of psychology that maps human consciousness and the evolution of human culture. The first idea we will cover is called the *Justification Hypothesis*, which provides an evolutionary account of human consciousness and culture and provides a clear framework for understanding why humans do not come equipped as abstract, analytic reasoners, but in fact operate primarily as socially motivated "reason givers." The second frame is the *Influence Matrix*, which is a map of the human relationship system, specifically the social motivations and emotions that intuitively guide people in relationships and social exchanges. Together, these two ideas capture the complex interplay between the social context and how humans justify their actions to themselves and to others, and provide us with a general lens from which to understand groupthink and related phenomena.

The Justification Hypothesis

The Justification Hypothesis (JH) is really several interlocking ideas bundled in one, and we will highlight two of its main features here. First, it is a theory of the origins of human self-consciousness, filling in the crucial missing link regarding the evolutionary forces that transformed the human mind into its modern form and gave rise to the explosion of human culture. The idea about the origins of human self-consciousness is what the "hypothesis" actually references. Second, the JH specifies the key domains of human consciousness and maps how they are interrelated via an updated "Tripartite Model."

The Justification Hypothesis on the Origins of Human Self-Consciousness

Human consciousness has always been the source of much fascination and mystery. Whether and how it emerged via evolution and natural selection has been controversial since the beginning of evolutionary models. Over the past several decades, a number of pieces of the evolutionary puzzle have come into view. Many scholars have highlighted how complex social interactions drive the evolution of higher intelligence in social mammals. The JH complements and adds to these perspectives via making an explicit connection between the modern design features of the human consciousness system and a novel evolved selection pressure that gave rise to it. Specifically, the JH points to how language changed the social environment, in particular by giving rise to a fundamentally new adaptive problem. The adaptive problem that emerged with a linguistic environment was the problem of social justification. The adaptive solution was an interpreter (see Gazzaniga, 1998) that provided justifications for actions that took social influence into account. Stated differently, as human language and cognitive capacities emerged, such that individuals began to ask questions that forced reason-giving accounts, the social psychological environment changed rapidly. This in turn shaped evolution of the human self-consciousness system into the "mental organ of justification" (Henriques, 2003).

Anyone who has raised a child knows that kids first learn simple commands and descriptions for objects (e.g., no, mommy, juice). After they obtain some mastery with descriptive language, a transformation happens, usually around the age of two; they start asking questions. The JH posits that the emergence of the "Q&A" capacity that tipped human evolution into a completely different phase. Why? Because, although asking questions is relatively easy, answering questions raises a completely new series of problems. To see what I mean, hang out with an intelligent, curious four-year-old who has discovered "why questions": "Why don't we eat cookies before we eat dinner?"; "Why are you going bald?"; "Why is the sky blue?" As such children readily demonstrate, asking questions is much easier than answering them. That is why exasperated parents eventually say, "That is just the way it is!"

Once language had tipped from descriptions and commands to a Q&A format, individuals could be held to account for why they did believe and act the way that they did. Although a chimpanzee can clearly send the message she is angry or scared, without a symbolic language it is almost impossible for her to communicate the reasons why she feels that way. In contrast, a "Q&A" language environment means humans can ask and be asked about the thought processes associated with their behaviors. Questions such as "Why did you do that?"; "What gives you the right to behave that way?" and "Why should I trust you?" force the issue. Obtaining information about what others think, what they have done, and what they plan to do is obviously important for navigating the social environment in modern times and, given that humans have always been an intensely social species, there is every reason to believe that it was equally essential in the ancestral past. The first basic claim of the JH is that once people developed the capacity to use language to access the thought patterns of others, they likely did so with vigor.

Now consider why the answers to those questions would have been so important. If you strike a comrade with a stick, it matters whether you tell him it was done by accident or on purpose. If you take more than your proportional share of meat, it matters how you explain that action. If you are bargaining with a stranger, you can get more resources if you emphasize why the resources you are trading are valuable, and so on. A second basic claim is that the kinds of explanations people offer for their behavior have real-world consequences.

A third claim is that human interests diverge and this complicates the interpretation process significantly. If one's interests always fully coincided with the interests of others, communicating the reasons for one's behavior would primarily be a technical problem of translating one's nonverbal thoughts into a symbolic form that could be understood. But because interests always diverge to some extent and the explanations given for one's behavior have real-world consequences, the communication task becomes one of justification rather than simple translation. These claims about the problem of social justification point directly to the design features we would expect the human self-consciousness system to exhibit. That is, to the extent that the adaptive problem of justification can be thought of as a "lock," the human self-consciousness system should look like a "key" that fits it.

So according to the Justification Hypothesis, what is the self-consciousness system? It is the language-based portion of your mind that is narrating what is happening, why it is happening, and why you are doing what you are doing in that context. This formulation clearly predicts that the self-consciousness system should be designed in such a way that it allows humans to effectively justify their actions to others in a manner that, all things being equal, tends to maximize social influence. An examination of some of the characteristics of human self-consciousness, as elucidated by neuropsychology, social, cognitive, and developmental psychology, demonstrates that there is a large body of general human psychological research that is highly consistent with this proposition. Specifically, Henriques (2011) reviews how the JH accounts for the "interpreter" function of the left hemisphere, cognitive dissonance, self-serving biases, motivational reasoning biases, findings on the differences between implicit and explicit attitudes, and research on reason-giving and

social accounts. In sum, the JH allows us to get the correct frame on the evolution of the self-consciousness system, and in so doing allows us to map the key domains of human consciousness and their interrelationship. This sets the stage for us to tackle a closely related concept, that of the human self.

The Updated Tripartite Model of Human Consciousness

The value of the Justification Hypothesis is not simply in that it provides a plausible story for why the human self-consciousness system might have evolved, but rather it sheds new light on understanding human consciousness. For example, the JH gives rise to an "updated" Tripartite Model of human consciousness (Fig. 2.1). It is referred to as an updated Tripartite Model because it divides human consciousness into three domains (the experiential system, the private self-consciousness system, and the public self-consciousness system) that parallels Freud's famous structural model in some regards (Henriques, 2003), although there are also crucial differences.

The experiential portion of human consciousness is quite different from the id that Freud envisioned. It is not an unconscious caldron of sex and aggression. Instead, it is a primary process, experiential system (see Epstein, 1994). This means
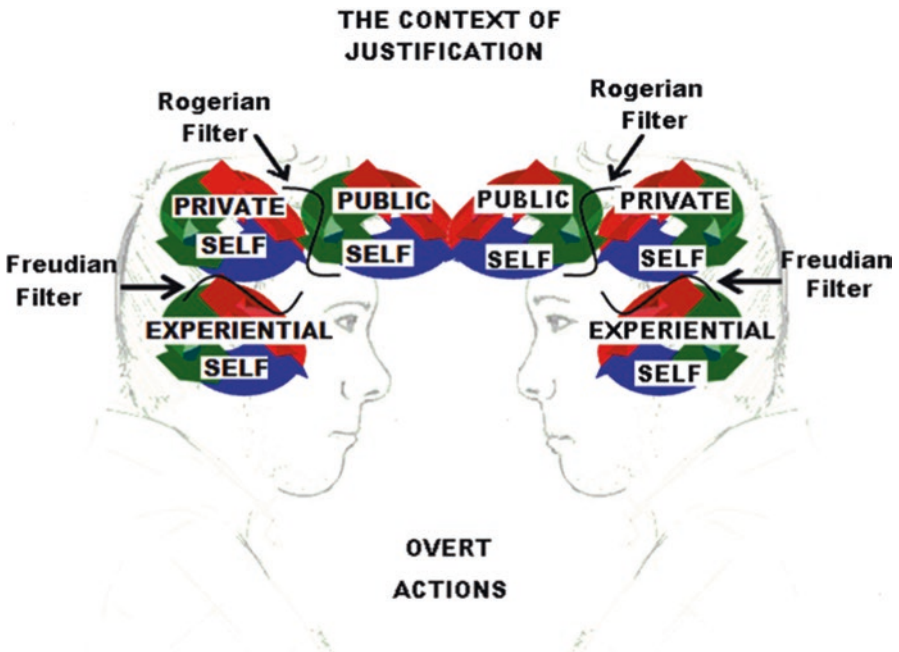

Fig. 2.1 The Updated Tripartite Model of Human Consciousness

that it is fast and relatively automatic. Everyday perception does not require conscious effort. Rather, you simply open your eyes and are presented with the world around you. The structure of the experiential system is framed in part by the Influence Matrix described in the subsequent section. As such, we can note here that the experiential system is part of the larger mind that is guided by positive and negative feeling states that function to orient an individual toward valued goals and away from dangers.

The private narrator portion shares some overlap with Freud's concept of an ego. However, Freud's conception of the ego was fuzzy. The JH achieves clarity because it comes with an explicit evolutionary account of the origin of the private self-consciousness system that helps us understand its design features, as described previously. The private narrator is a reflective reason-giving system that helps to navigate the relational world. It is clearly not fully formed at birth, but instead develops with experience. It requires the cognitive capacity to become the object of one's own attention. It also requires the capacity for symbolic language because it is fundamentally a reason-giving structure. In young children, the system operates to learn the rules of conduct. In adolescence, it emerges as a potentially separable identity, such that the teenager can reflect on the distinction between how they are actually and how they wished they might be. Finally, in adults it becomes a full narrator, an active self-concept that is weaving together one's life story (see McAdams, 2013).

With its central focus on justification and reason giving in a social context, the JH highlights a crucial distinction between private and public domains of justification. Anyone who has accidently shared a thought or action publicly that was meant to be private will quickly and powerfully experience the distinction. For example, on an e-mail listserv, I once accidently sent a message that was meant to be back channel to a single individual and not to the entire group. As soon as I hit send and saw the message appear on the board for everyone to see, a jolt of fear and anxiety rippled through me, and I literally let out a yelp. The distinction of private and public becomes very salient when barriers that are supposed to function to separate the two fail.

One of the most powerful pieces of evidence for the JH is the way it characterizes the relationship between the three domains of consciousness. Specifically, it highlights the presence of filters between the experiential and the private, and the private and the public. These are the "Freudian" and "Rogerian" filters, respectively, and they are clearly framed by the logic of the JH, which is why they are explicitly labeled in the diagram. The Freudian filter (or experiential-to-private filter) works via the process of inhibiting disruptive or problematic feelings, images, or thoughts, and shifting attention away from them. Why are certain impulses filtered? According to the Justification Hypothesis, the reason is to maintain a consistent, relatively stable justification narrative of the self and to maintain a justifiable image in the eyes of others. In his book *Ego Defenses and the Legitimization of Behavior*, Swanson (1988) made exactly this point, explicitly arguing that we should think of all ego defenses as "justifications that people make to themselves and others—justifications so designed that the defender, not just other people, can accept them" (p. 159). Part of the filter also involves a shift in attention, which can be called repression or, in behavioral terms, "experiential avoidance." Such experiential

avoidance is often supported by justifications, such as "There is no point in my feeling sad about that, it only brings up pain."

The Rogerian (or private-to-public filter) refers to the extent to which we share or do not share our narrative with others. It also refers to the way we share such information. Any time that you are thinking about whether or not to share a piece of information with another person or group of people, this is an example of the private-to-public filter. The JH posits that we all have significant experience filtering our thoughts, depending on who the audience is and what we want them to see in us. The example I referenced regarding the e-mail accidently going out to the whole group demonstrates that this filtering process is not just one on one, but refers to how we navigate our identities to different audiences across the levels of human interaction, from self to dyad to small group to full public identity. When diaries are sealed or doors are closed or memos are marked "confidential," we can see clearly the private-to-public filter at work.

The three domains of consciousness are not the only aspects in Fig. 2.1. Above the two figures is labeled "The Context of Justification," which refers to the network of symbolically based beliefs and values that provide the interacting members a shared frame of reference for their interaction. The context can be considered on the dimensions of time and scope. Scope refers to the size and scale of the context one is considering. Urie Bronfenbrenner's (1979) Ecological Systems Theory provides a useful framework for considering the scope of the system, although I should note that he was concerned with the whole societal context, which would include the biophysical ecology and technology, in addition to the systems of justification. There is also the concept of action, which refers to the overt activities of each individual.

The Influence Matrix

The above description offers some details on the domains of justification (i.e., the private and the public) and argues that humans create and live in contexts of justification, but the model does not clearly articulate what exactly drives the system. That is, where does the energy come from to justify one's actions in a certain way? The unified theory of psychology answers this question in the form of *Behavioral Investment Theory* (BIT). The BIT posits that the "mind/brain system" is evolved as a computational control center that computes the animal's actions on an energy investment value system built via natural and behavioral selection (or evolution and learning), such that animals are inclined to move toward "the good" (which they seek and approach) and away from "the bad" (which they avoid and withdraw from). The Influence Matrix is an extension of BIT, applied to human social motivation and emotion.

Based on much research in *personality* and social psychology, *the Influence Matrix* (Fig. 2.2) offers a map of our foundational relational strivings. The Matrix identifies one core relational motive and then highlights several other key relational
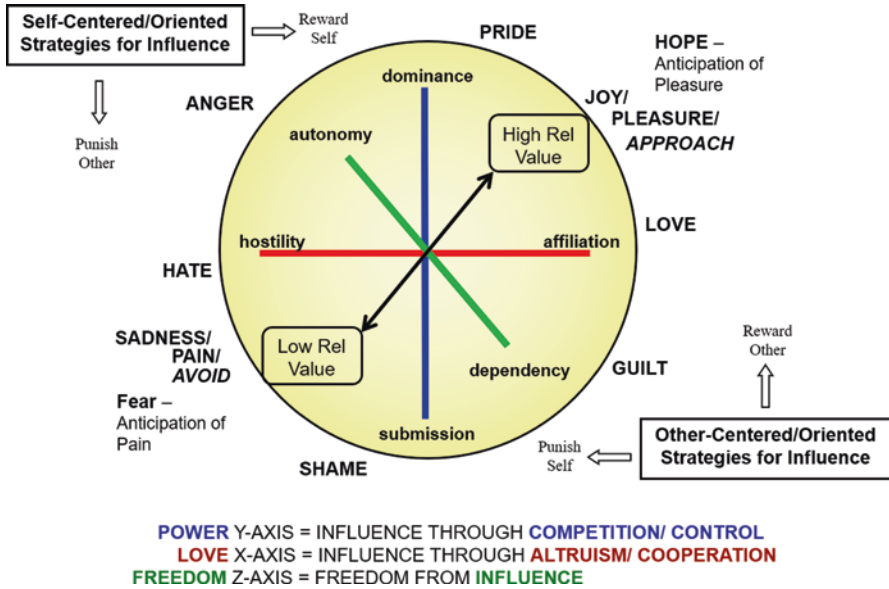
Fig. 2.2 The Influence Matrix

motives that are connected in various ways to the core motive. The core motive is called the Relational Value/Social Influence (RV-SI) motive, and is represented on the center diagonal line in the diagram, with high relational value on one end and low relational value on the other. This line refers primarily to the degree to which you feel known and valued by important others (i.e., family, friends, lovers, groups one belongs to). Behaviorally, it refers to the extent to which important others will share and invest in your interests, which is the degree of social influence one has – you put them together to get the RV-SI line. From this perspective, people are energized to be respected, valued, appreciated, cared for, or admired. And they are also energized to avoid the opposites (i.e., being rejected or held in contempt or ignored by important others) and we can see these are things folks generally fear and try to avoid. The vast majority of people can see that these are major themes in their relational world.

Although RV-SI is the core motive, we can go further and assert there are various ways folks try to gratify their RV-SI needs. There are two "competitive" (or vertical) ways folks try to obtain RV-SI. One common and directly competitive way is via power. *Power* in the form of direct dominance, *leadership*, and control over others is a way to insure social influence. Another competitive relational process, but one that is more indirect, is *achievement*. Achievement refers to accomplishing markers of skill and status which are valued in a society or group. These two motives (i.e., direct and indirect competitive influence) are represented by the vertical line.

There are also two forms of positive "horizontal" or cooperative relating to acquiring social influence and relational value. These motivational forces and kinds of relating involve affiliating and joining one's interests with the interests of others.

One form of affiliative motive is called belonging, which refers to being a part of a group or *identity*. So, if you take pleasure in rooting for a sports *team* or feel a close identification with your *religious* group or nationality, then you are experiencing the motive to belong. The other affiliative urge is intimacy. This involves letting others know more about "the real you" and joining with them at a much more personal level. Intimacy involves breaking down public filters and sharing authentically to allow for a more genuine joining of interests and private feelings.

We have identified a core relational motive (RV-SI) and four common relational strivings (achievement, power, belonging, and intimacy). Two other common relational strivings include independence and *self-reliance*. Although we are clearly social animals and deeply seek connection and approval, it is also the case that we are in need of separation and individuation to combat being completely dependent on the whims and desires and opinions of others. To the extent that an individual advertises one's self-reliance, diminishes their needs for approval and connection, or tries to buck the trend, defy social norms, and carve their own unique path, they are engaged in autonomous strivings.

Finally, there are relational "avoidance strategies," where individuals strive to avoid the negative consequences of trying to achieve the approach strivings. Submitting or surrendering in *competition* is one such avoidance strategy. Many individuals are plagued by self-conscious, shameful thoughts about how inferior they are. The root of this behavior is that these folks are striving to avoid competition or conflict which would then cause them to lose respect or be embarrassed. Whereas *shame* and submission are about avoiding relational conflict and competition, hostility and contempt are about avoiding affiliation or connection (and often then justifying power). These are "othering" strategies designed to avoid betrayal and others' control and to remove any sense of obligation to them (i.e., make them unimportant). In existing affiliative relations, we use *anger* and hostility to remind those close to us of their obligations to us and to remove or diminish their tendency to betray us (or we use it to move away from them after we feel they have betrayed us and we can no longer have an intimate relation with them). Notice that a difference between avoidance strategies and approach strategies is that folks do not engage in avoidance strategies just for the sake of doing so. Very few people strive to be hostile or ridden with shame. But they are activated in the service of avoiding some even worse outcome.

Implications for Understanding Groupthink and Other Related Phenomena

Why do people believe what they do? How do social dynamics influence how people reason? Why would a group of people, who say averaged a left leaning "8" on a political scale ranging from 1 to 10, come up with a political "solution" that would be rated an "11" by individuals external to the group? The empirical elements that

contribute to groupthink are well documented. When there is pressure on a group to believe a certain way, when there is a strong need for unity, often spurred by the identification of an "other" group that is seen as a polar opposite, when emotions are stoked, and when there is a charismatic leader who wants to see certain things done, the stage is set for groupthink. What these empirical findings do not answer are ultimate questions about why humans believe, feel, and act in this way. They do not provide models regarding the architecture of the human mind or its motivational systems that clearly delineate why these processes unfold.

It is this latter gap that the combination of the Influence Matrix and Justification Hypothesis fills. When taken together, these models state very clearly that human reasoning is not a cold analytic process, designed to take in information and calculate pros and cons via some "rational actor." Rather, human consciousness is guided by the need for social influence and relational value, and more specific needs for power and achievement or belonging and intimacy or the avoidance of the loss of such things. And human consciousness functions, first and foremost, as a social reason-giving system, one that seeks a personally and publicly socially justifiable path to legitimize action. If we understand this as the fundamental model by which humans operate, then the phenomena of groupthink, along with many other social psychological processes, become readily understandable.

References

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709–724.

Gazzaniga, M. S. (1998). *The mind's past*. Berkeley, CA: University of California Press.

Henriques, G. R. (2003). The tree of knowledge system and the theoretical unification of psychology. *Review of General Psychology, 7*, 150–182.

Henriques, G. R. (2011). *A new unified theory of psychology*. New York: Springer.

Henriques, G. R. (2017, August). A moral compass for troubling times. *Theory of Knowledge on Psychology today*. Retrieved February 14, 2019, from https://www.psychologytoday.com/us/blog/theory-knowledge/201708/moral-compass-troubling-times.

Lilla, M. (2017). *The once and future liberal: After identity politics*. New York: Harper Collins.

McAdams, D. P. (2013). The psychological self as actor, agent, and author. *Perspectives on Psychological Science, 8*, 272–295.

Swanson, G. E. (1988). *Ego defenses and the legitimation of behavior*. New York: Cambridge University Press.

Turner, M. E., & Pratkanis, A. R. (1998). Twenty-five years of groupthink theory and research: Lessons from the evaluation of a theory. *Organizational Behavior and Human Decision Processes, 73*(2–3), 105–115. https://doi.org/10.1006/obhd.1998.2756