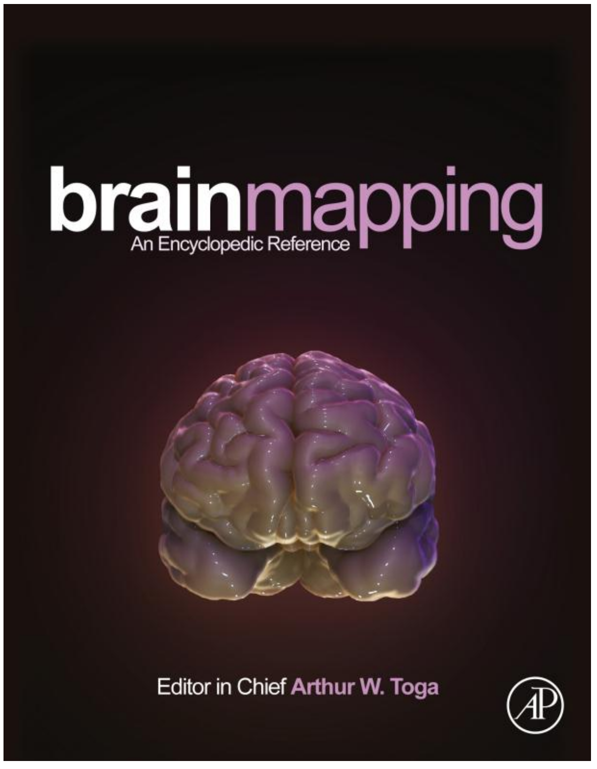

Spunt R.P. (2015) Dual-Process Theories in Social Cognitive Neuroscience. In: Arthur W. Toga, editor. *Brain Mapping: An Encyclopedic Reference*, vol. 3, pp. 211-215. Academic Press: Elsevier.

Dual-Process Theories in Social Cognitive Neuroscience

RP Spunt, California Institute of Technology, Pasadena, CA, USA

Glossary

Awareness The extent to which a subject has conscious access to the stimulus that initiates a process, the operation of the process itself, and/or the output of the process.
Controllability The extent to which the course of a process can be altered (i.e., modulated or terminated) after it has begun.
Efficiency The extent to which a process can be executed quickly and in the absence of attention.
Intentionality The extent to the initiation of a process is subject to voluntary.

Introduction

Social cognition broadly refers to the set of cognitive abilities used to understand and interact with conspecifics. In humans, such abilities are remarkably advanced, enabling a level of sociality that far outstrips that which is observed in even our closest primate ancestors (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007). Social psychologists have long recognized that the complexity, frequency, and uncertainty of social information necessitate parallel information-processing streams that are automatically initiated by the presence of relevant stimuli and are executed without requiring intentional control or conscious deliberation (Bargh & Chartrand, 1999; Smith & Decoster, 2000). This distinction between automatic processing and controlled processing is foundational to a family of theories known as *dual-process theories* (Chaiken & Trope, 1999; Sherman, Gawronski, & Trope, 2014).

Today, many social psychologists are using a social cognitive neuroscience approach to test theories about social cognition (Lieberman, 2010). Such an approach has used neuroimaging methods such as functional magnetic resonance imaging (fMRI) to measure social information processing as it unfolds in the brain. Neuroimaging data have been primarily used for the purposes of localizing specific social cognitive processes to specific regions or networks in the brain. By investigating how these regions function under different conditions, social psychologists are now in a position to utilize neuroimaging measures to evaluate and test dual-process theories of the social mind. Surprisingly, dual-process theories have received limited application in the design, analysis, and interpretation of neuroimaging data (but see Satpute & Lieberman, 2006; Spunt & Lieberman, 2014). In this article, I aim to facilitate such an application by providing a comparison of categorical to dimensional dual-process frameworks for conceiving the operation of the processes that collectively make up our social abilities. Both frameworks assume that the functional role of any cognitive process can be described in a manner that is independent of its operating characteristics, of which we follow Bargh (1994) in including four: consciousness, intentionality, controllability, and efficiency.

The Categorical Framework

Figure 1(a) illustrates the categorical framework, so-named because it treats automatic and controlled processing as distinct processing categories: Automatic processes are those whose operation is unconscious, unintentional, uncontrollable, and efficient, while controlled processes are those whose operation is conscious, intentional, controllable, and inefficient.

Hence, the categorical view assumes (sometimes only implicitly) perfect covariance among these four operation dimensions. Based on this assumption, any given process can be categorized as automatic or controlled by assessing its position on any one of the four operating dimensions identified in the preceding text (Gawronski, Sherman, & Trope, 2014). For example, a process that has been shown to operate unconsciously (e.g., using subliminal priming methods, see Bargh & Chartrand, 2000) would be also assumed to also operate unintentionally, uncontrollably, and efficiently.

Most categorical dual-process theories are also dual-system in that they specify the existence of different cognitive systems for implementing automatic versus controlled processes (Bargh & Chartrand, 1999; Chaiken & Trope, 1999; Epstein, 1994; Evans, 2008, 2010; Gilbert, 1989; Kahneman, 2003; Schneider & Shiffrin, 1977; Sherman et al., 2014; Sloman, 1996; Smith & Decoster, 2000). For example, Stanovich and West (2000) used the generic labels system 1 and system 2 to refer to automatic and controlled systems, respectively. Figure 1(b) displays a fictional illustration of what is implied by a dual-system view of brain function: Some regions are categorized as automatic, and some regions are categorized as controlled. In the brief history of social cognitive neuroscience, the most ambitious attempt to conceive brain function categorically is the X-system and C-system model proposed by Lieberman, Gaunt, Gilbert, and Trope (2002).

It is becoming increasingly clear that the categorical framework, although intuitive and in many cases useful, glosses over a great deal of complexity (Bargh, 1989, 1994; Gawronski et al., 2014; Moors & De Houwer, 2006; Spunt & Lieberman, 2014). The most compelling critique of categorical dual-process theories regards the assumption of perfect covariance among the four dimensions. Numerous behavioral studies suggest that these dimensions are to some extent orthogonal and, correspondingly, that different methods are required to test hypotheses about different dimensions.

The Dimensional Framework

An alternative to the categorical framework is to conceive and investigate each dimension in its own right. Hence, in this

Brain Mapping: An Encyclopedic Reference, (2015), vol. 3, pp. 211-215

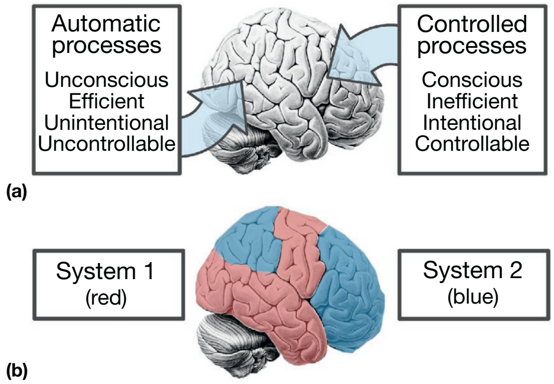

Figure 1 (a) Illustration of the categorical view of automatic and controlled processes as applied to brain function. The positions of the arrows illustrate the conflation of bottom-up processing with automaticity, and top-down processing with control. (b) Illustration of what is implied by a dual-system view, in which regions of the brain are assigned to one of two systems.

section, I will define each dimension separately and will consider how they might be investigated using neuroimaging methods.

Awareness

Awareness refers to the extent to which the subject is phenomenally aware of the stimulus that initiates a process, the operation of the process itself, and/or the output of the process. What does it mean to say that a person has awareness of a neural process? In the current state of the psychological and brain sciences, this is more of a metaphysical question than a scientific one. However, from a methodological perspective, one can establish operation without awareness by testing if subliminal stimulus variation impacts neural activity. In terms of experimental design, the logic is identical to behavioral studies employing subliminal primes (Bargh & Chartrand, 2000). Yet unlike behavioral studies, the researcher is not dependent on measuring a behavioral outcome (e.g., response latency or recall) to establish that the subliminal stimulus induced a nonconscious process. Instead, measures of brain activity can be used to establish nonconscious processing. For example, numerous fMRI studies have now observed amygdala sensitivity to variation in the emotional expression (Whalen et al., 1998) or value (Morris, Ohman, & Dolan, 1999) of faces presented subliminally. On the basis of such data, one can conclude that some component of face processing in the amygdala can occur in the absence of awareness.

Another method for tapping process awareness involves establishing a disconnect between self-reported behavioral intentions and a neural process known to be associated with the intended behavior. To the extent that such a disconnect is observed, one can conclude that the subject is unaware of the neural process. Using this logic, researchers have suggested that the formation of movement intentions operates unconsciously insofar as neural activity associated with the production of movements actually precedes participants' self-reported decision to move (Desmurget & Sirigu, 2009). Using a similar logic, work from our group has shown in several studies that neural processes at work during the consumption of persuasive messages do a better job of predicting participants' ensuing behavior than does the participants' own self-reported intentions to engage in those behaviors (Falk, Berkman, Mann, Harrison, & Lieberman, 2010; Falk, Berkman, Whalen, & Lieberman, 2011). On the flip side, claims that a neural process is available to awareness can be supported by observations of tight coupling between self-reported phenomenal states (e.g., frustration) and a neural process (e.g., the neural response to errors; Spunt, Lieberman, Cohen, & Eisenberger, 2012). These studies suggest that, in addition to the use of subliminal stimulus variation, the careful assessment of (self-reported) phenomenal experience can be used to make claims about processing (un)awareness.

Efficiency

The efficiency of a neurocognitive process can be defined as the extent to which it can be executed quickly and in the absence of attention. Phenomenologically, efficient processes are fast and effortless (e.g., recognizing a familiar face), while less efficient processes are slow and effortful (e.g., mentally computing the product of 42 and 79). Importantly, efficient processing is not the same as spontaneous processing. Spontaneity simply means that a process will be engaged even in the absence of external stimulus to do so (e.g., the instruction to multiply 42 and 79). Hence, a process that begins spontaneously need not operate efficiently.

The principal method for studying processing efficiency is the dual-task paradigm in which the researcher gives the subject a primary task known to engage the process of interest and simultaneously varies either the presence or difficulty of a secondary task (often called a manipulation of 'cognitive load'). When a dual-task paradigm is employed in a behavioral study, questions about efficiency can only be answered if both tasks produce measurable behavioral outcomes that validly reflect successful engagement of the primary and secondary processes. When such outcomes are available, one can confirm that cognitive load was successfully induced (by examining performance of the secondary task) and assess the extent to which it affects performance on the primary task. If performance on the primary task is unaffected by load, one can infer that the process associated with the primary task is efficient.

Dual-task paradigms can be used to study neural efficiency by engaging the neural process of interest in a primary task while simultaneously varying either the presence or difficulty of a secondary task. However, when measuring brain activity, one is no longer dependent on observations of performance variability, since variability in the neural response is now a (more) direct measure of the process of interest. Hence, primary tasks that do not produce a behavioral outcome, such as supraliminal stimulus manipulations, can be used to study neural efficiency (Pessoa, Mckenna, Gutierrez, & Ungerleider, 2002). However, it should be emphasized that in order for a supraliminal stimulus manipulation to provide evidence of neural efficiency, it must be paired with a cognitively demanding secondary task so that strategic stimulus processing can be ruled out as an alternative explanation. Moreover, we note that to categorize such a paradigm as 'dual-task' is perhaps a misnomer, since passive stimulus perception is not a 'task' in the traditional sense. Hence, it may be more appropriate to characterize this type paradigm as involving 'task-independent stimulus variation.'

The use of dual-task manipulations to test dual-process theories of social cognition is illustrated in Spunt and Lieberman (2013). This study used a traditional cognitive load manipulation (digit retention) with fMRI to investigate the efficiency of attributional processing during action observation. When under minimal load, attributional (relative to factual) processing of another person's actions evoked activity in two regions that previous studies have reliably associated with making attributions for actions (Spunt & Lieberman, 2012; Spunt, Satpute, & Lieberman, 2011); when load was increased, this effect was extinguished as predicted by the dual-process theories of Trope (1986) and Gilbert (1989), both of which propose that key steps of the attributional process are inefficient in that they depend on the presence of sufficient attentional resources. In contrast, activity in regions believed to support the recognition of motor actions was largely insensitive to the load manipulation, a finding predicted by theories of the motor system's role in social cognition (Gallese, 1998).

Intentionality

The intentionality of a mental process regards the extent to which its initiation depends on the presence of an explicit intention to initiate it. Methodologically, we consider two ways in which the intentionality of a region's operation can be investigated. The first involves manipulating the subject's goal while presenting no stimulation. Using such a method, researchers have, for example, demonstrated that visual cortices of the brain are subject to intentional control insofar as they activate when subjects intentionally engage in visual imagery (Kosslyn, Ganis, & Thompson, 2001). A second method involves manipulating the subject's goal while either holding stimulation constant (Spunt, Falk, & Lieberman, 2010) or crossing the goal manipulation with a stimulus manipulation (Spunt & Lieberman, 2012; Winston, Strange, O'doherty, & Dolan, 2002).

Controllability

Controllability refers to the extent to which the course of a process can be altered (i.e., modulated or terminated) after it has been initiated. To investigate controllability, the researcher must first induce the process of interest (e.g., a negative emotional response to an aversive stimulus) and then manipulate the presence of a goal to in some way alter the process (e.g., diminish the negative emotional response by reappraising the stimulus; Ochsner et al., 2004). This method allows the researcher to assess the dynamic interaction of bottom-up (i.e., stimulus-driven) and top-down (i.e., goal-driven) psychological processes. In order to do so, studies employing this method typically proceed by identifying a region (or set of regions) as implementing the top-down process (e.g., searching the brain for regions associated with the presence and/or successful execution of the control goal) and a region (or set of regions) whose association with the bottom-up process is affected by the presence and/or successful execution of the control goal. Furthermore, these studies can employ a variety of connectivity analyses to determine if activation of the region identified as 'controller' exhibits dynamic changes in activity that track goal-dependent changes in activation of the region (s) identified as 'target' (Lieberman, 2011). When the process to be controlled is initiated by presentation of a stimulus, researchers should be careful in interpreting goal-dependent changes in neural processing of the stimulus (Pessoa et al., 2002). This is because these changes could be caused by a direct influence of the controller region(s) on the target region(s) or by an indirect path in which the controller region(s) produces shifts in spatial attention that ultimately alter the bottom-up inputs to target regions (Posner, 1980). This issue can be at least partially addressed by tracking eye movements (Dalton et al., 2005).

It has been suggested that the central component of control is working memory, which allows for online maintenance and serial manipulation of mental representations. Hence, one method for establishing the controllability of a process is to determine whether it operates like a working memory system (Evans, 2008). The most common method for investigating working memory is parametrically manipulating the amount of information to be held in mind. However, studies of working memory have almost exclusively investigated the maintenance and manipulation of nonsocial information, such as numbers and sensorimotor representations. These studies reliably find that a lateral frontoparietal network shows increases in activity that parametrically track increasing amounts of information load (Cabeza & Nyberg, 2000). We developed a paradigm for investigating the online maintenance and manipulation of social representations, that is, social working memory (Meyer, Spunt, Berkman, Taylor, & Lieberman, 2012). While undergoing fMRI, participants were asked to rank two, three, or four of their friends on a trait dimension (e.g., generosity) during a 6 s delay period. At the end of the period, they made a rank judgment (e.g., Is Rebecca the second most generous?) that could be coded as accurate or inaccurate based on ratings collected from the participant prior to the fMRI study. When they accurately ranked their friends, regions associated with making mental state inferences in prior studies demonstrated parametric increases in activity as a function of the number of friends to be ranked. This strongly suggests that these regions can be intentionally controlled in a top-down fashion to manage increasing social cognitive demands.

Conclusion

In conclusion, evidence from the social cognitive neurosciences provides little evidence for a categorical view of cognitive processing in the human brain. Instead, such processing is best characterized in terms of a set of potentially orthogonal operating dimensions: awareness, efficiency, intentionality, and controllability. This multidimensional view of cognitive processing is becoming the default view of information processing in social psychology (see Sherman et al., 2014). Hence, instead of making claims that categorize a neurocognitive process as either automatic or controlled, such claims should position a process across a multidimensional space that might be termed the processes' *automaticity profile* (Spunt & Lieberman, 2014).

It would be misleading for such claims to say that specific regions of the brain possess specific automaticity profiles. This is because a neural process is not only multidimensional but also highly dependent on the conditions under which it is called upon to operate. This is powerfully illustrated with reference to the amygdala, which is often invoked as a paradigmatic example of neural automaticity (Dolan & Vuilleumier, 2003; Ohman, 2002; Satpute & Lieberman, 2006). Although studies have shown that this region responds to subliminally presented threat cues (e.g., Morris et al., 1999), other work suggests that the automatic response of the amygdala to such cues is dependent on the presence of sufficient attention to the sensory modality in which such cues are presented (Pessoa et al., 2002). In a recently published study, Mothes-Lasch, Mentzel, Miltner, and Straube (2011) simultaneously presented subjects with threatening auditory cues (angry or neutral voices) and neutral visual stimuli (cross or circle) and manipulated whether they judged the gender of the speaker or the type of visual symbol. Replicating previous research, the amygdala robustly responded when participants attended to the voice during the gender judgment task. However, this effect was extinguished when attention was instead directed toward the visual modality in the symbol judgment task. Studies such as this one suggest a conditional view of automaticity in the amygdala that can be phrased as follows: when a threat cue is present in a sensory modality to which attention is being directed, the amygdala can detect it even in the absence of explicit awareness.

Thus, in addition to a multidimensional approach presented earlier, it is important to recognize that neural automaticity is conditional (Bargh, 1989; Spunt & Lieberman, 2014). This is because regions do not operate in isolation but are part and parcel of distributed functional networks that are constantly being modulated by changes in the internal and external environment (Pessoa, 2008; Pessoa & Adolphs, 2010). Even holding the region constant, function can vary dramatically depending on the conditions under which it is engaged (Poldrack, 2006). Overall, this analysis emphasizes the necessity of conditional statements regarding the automaticity of a brain region or network. When such conditional statements are combined with a region's position in multidimensional operating space, its function can then be described in a manner that is more consistent with what is known about how the brain carries out the work of social cognition.

See also: **INTRODUCTION TO COGNITIVE NEUROSCIENCE**: Response Inhibition; Top-Down Suppression; **INTRODUCTION TO SOCIAL COGNITIVE NEUROSCIENCE**: Action Perception and the Decoding of Complex Behavior; Emotion Regulation; Mentalizing; Strategic Mentalizing: The Neural Correlates of Strategic Choice; The Amygdala and Social Perception.

References

Bargh, J. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Hillsdale: Erlbaum.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*(7), 462–479.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). New York: Cambridge University Press.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: The Guilford Press.

Dalton, K., Nacewicz, B., Johnstone, T., Schaefer, H., Gernsbacher, M., & Goldsmith, H. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*(4), 519–526.

Desmurget, M., & Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences*, *13*(10), 411–419.

Dolan, R. J., & Vuilleumier, P. (2003). Amygdala automaticity in emotional processing. *Annals of the New York Academy of Sciences*, *985*(1), 348–355.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *16*.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Psychology*, *59*, 255–278.

Evans, J. S. B. (2010). *Thinking twice: Two minds in one brain* (1st ed.). Oxford University Press.

Falk, E., Berkman, E., Mann, T., Harrison, B., & Lieberman, M. (2010). Predicting persuasion-induced behavior change from the brain. *Journal of Neuroscience*, *30*(25), 8421–8424.

Falk, E., Berkman, E., Whalen, D., & Lieberman, M. (2011). Neural activity during health messaging predicts reductions in smoking above and beyond self-report. *Health Psychology*, *30*(2), 177–185.

Gallese, V. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501.

Gawronski, B., Sherman, J. W., & Trope, Y. (2014). Two of what? A conceptual analysis of dual-process theories. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 3–19). New York, NY: Guilford Press.

Gilbert, D.T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*(5843), 1360–1366.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720.

Kosslyn, S., Ganis, G., & Thompson, W. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635–642.

Lieberman, M. (2010). Social cognitive neuroscience. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 143–193). New York: McGraw-Hill.

Lieberman, M. (2011). Why symbolic processing of affect can disrupt negative affect: Social cognitive and affective neuroscience investigations. In A. Todorov, S. T. Fiske, & D. Prentice (Eds.) (pp. 188–209). New York: Oxford University Press, http://dx.doi.org/10.1093/acprof:oso/9780195316872.003.0013.

Lieberman, M., Gaunt, R., Gilbert, D., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology*, *34*, 199–249.

Meyer, M., Spunt, R., Berkman, E., Taylor, S., & Lieberman, M. (2012). Evidence for social working memory from a parametric functional mri study. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(6), 1883–1888.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326.

Morris, J., Ohman, A., & Dolan, R. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(4), 1680–1685.

Mothes-Lasch, M., Mentzel, H., Miltner, W., & Straube, T. (2011). Visual attention modulates brain activation to angry voices. *Journal of Neuroscience*, *31*(26), 9594–9598.

Ochsner, K., Knierim, K., Ludlow, D., Hanelin, J., Ramachandran, T., Glover, G., et al. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, *16*(10), 1746–1772.

Ohman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science*, *11*(2), 62–66.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158.

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–783.

Pessoa, L., Mckenna, M., Gutierrez, E., & Ungerleider, L. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(17), 11458–11463.

Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Satpute, A., & Lieberman, M. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Research*, *1079*(1), 86–97.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Sherman, J. W., Gawronski, B., & Trope, Y. (Eds.). (2014). *Dual-process theories of the social mind*. New York: The Guilford Press.

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22.

Smith, E. R., & Decoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108–131.

Spunt, R., Falk, E., & Lieberman, M. (2010). Dissociable neural systems support retrieval of how and why action knowledge. *Psychological Science*, *21*(11), 1593–1598.

Spunt, R., & Lieberman, M. (2012). Dissociating modality-specific and supramodal neural systems for action understanding. *Journal of Neuroscience*, *32*(10), 3575–3583.

Spunt, R., & Lieberman, M. (2013). The busy social brain: Evidence for automaticity and control in the neural systems supporting social cognition and action understanding. *Psychological Science*, *24*(1), 80–86.

Spunt, R. P., & Lieberman, M. D. (2014). Automaticity, control, and the social brain. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (2nd ed.). New York: Guilford Press.

Spunt, R., Lieberman, M., Cohen, J., & Eisenberger, N. (2012). The phenomenology of error processing: The dorsal ACC response to stop-signal errors tracks reports of negative affect. *Journal of Cognitive Neuroscience*, *24*(8), 1753–1765.

Spunt, R., Satpute, A., & Lieberman, M. (2011). Identifying the what, why, and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, *23*(1), 63–74.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645–665.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*(3), 239–257.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., Mcinerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience*, *18*(1), 411–418.

Winston, J., Strange, B., O'doherty, J., & Dolan, R. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*(3), 277–283.