Lost in dissociation:

The main paradigms in unconscious cognition.

Luis M. Augusto

Independent scientist / University of Barcelona (visiting researcher) [at completion of paper]

Spain

Published in Consciousness and Cognition, 42 (2016), 293-310

Abstract: Contemporary studies in unconscious cognition are essentially founded on dissociation, i.e., on how it dissociates with respect to conscious mental processes and representations. This is claimed to be in so many and diverse ways that one is often lost in dissociation. In order to reduce this state of confusion we here carry out two major tasks: based on the central distinction between cognitive processes and representations, we identify and isolate the main dissociation paradigms; we then critically analyze their key tenets and reported findings.

Keywords: Unconscious cognition; the dissociation logic; exclusiveness assumption; exhaustiveness assumption; cognitive formats or levels; representation paradigms; processing paradigms

1. Introduction

We have long suspected that there is (far) more to our mental life than we are conscious of (see, e.g., Ellenberger, 1970). Evidence abounds that points in this direction: for example, we often reach conclusions and find solutions to problems without actually being aware of the reasoning processes, and we frequently cannot tell what knowledge basis we draw from when making important decisions and judgments, such as judging faces or investing our hard-won money. Today, we have the scientific means to approach this hypothesis, and investigation into this field now permeates the whole of psychology and cognitive science. In fact, more than sufficient contributions are today available to constitute a sub-discipline of unconscious cognition (see Augusto, 2010).

The study of unconscious cognition has grown in importance and more and more fields became interested in its many facets, such as knowledge management (e.g., Bennet & Bennet, 2008), education (e.g., ECU, 2013; Sun, Mathews, & Lane, 2007), medical care (e.g., Sabin & Greenwald, 2012), consumer behavior (e.g., Dijksterhuis et al., 2005), the law (e.g., Casey et al., 2012), artificial intelligence (e.g., Schank, 2009), and even finance (e.g., Taffler & Tuckett, 2010). However, it has also become more complex, namely with respect to dissociation, of which there is today a plethora of paradigms.

In practice, in an obvious way, *dissociation* appears to be the right approach to study both consciousness and unconscious cognition (*the dissociation logic*; Timmermans & Cleeremans, 2015). It is largely at the basis of the highly influential *dual-process/system theories* that postulate an architectural—and often evolutionary—dissociation between the overall unconscious System 1 and the conscious processing of

System 2 (Evans, 2008; Kahneman & Frederick, 2002; Stanovich, 1999; see also Sherman et al., 2014).

The fact of the matter is that those approaching the field of unconscious cognition are often "lost in dissociation," because a unifying theoretical and methodological framework is largely missing and there reigns terminological and conceptual confusion. The very concept of dissociation has motivated some important caveats (see Augusto, 2013), and criticism has fallen on the double-process/system theories (e.g., Evans, 2008; Evans & Stanovich, 2013), as well as on the criteria, thresholds, measures and testing methods in unconscious cognition (e.g., Fisk & Haase, 2005; Reingold, 2004; Timmermans & Cleeremans, 2015; see Section 2.2). Under closer scrutiny, many claimed dissociations lose their clean-cut character, sometimes to the point of effacement. Briefly, confusion still reigns.

In order to alter this state of affairs we here carry out a double major task: based on the central distinction between cognitive processes and representations, we identify and isolate the main dissociation paradigms, and we proceed to analyze them critically with respect to their key tenets and reported findings. While the former task is expected to contribute both to settle some major terminological and conceptual issues and to reduce the number of dissociations in the literature, the latter will hopefully motivate a more cautious usage of the claimed dissociations, at least until further work is carried out addressing the main issues here exposed.

2. Dissociating conscious from unconscious cognition

2.1. The notion of dissociation

As would be expected, after centuries of a deeply rooted philosophical tradition equating cognition with consciousness (e.g., Brentano, 1874/1973; Descartes, 1644/1983; Locke, 1690/1959), the scientific hypothesis that humans often—or more often than not—think without being aware of their own thoughts, and often without being capable of becoming so, profoundly agitated western cultures at the end of the 19th century. Nevertheless, the idea took root, and the notion that there was a *dissociation* or *splitting of consciousness* making for a *double consciousness* or *self*, shared by psychopathologists and early experimental psychologists alike (e.g., Binet, 1890; Breuer & Freud, 1895; Sidis, 1898), became pervasive and influential in early 20th-century psychology and psychiatry.

To be sure, not many today still speak of *doubles*, but most contemporary research in the field of consciousness is grounded on the presupposition that consciousness and awareness must be investigated by following theoretical guidelines and by applying experimental methods that might reveal how conscious and unconscious cognition *dissociate* (e.g., Timmermans & Cleeremans, 2015; more generally: Overgaard, 2015).

Paradigms in unconscious cognition all aim primarily at *falsifying the null model*, i.e., the model according to which there is no unconscious cognition. The general aim is to show that when one analyzes cognitive output from experimental subjects in face of a given input, there is far more to it than they are aware of or than is directly observable in behavioral terms. Unconscious cognition is thus invoked to account for this mismatch and conscious and unconscious cognition are said to dissociate.

The notion of dissociation here at play is to some extent different from the *dissociation of function* that lies in the foundations of (cognitive) neuropsychology (e.g., Coltheart, 2001). In the field of unconscious cognition, one speaks of dissociation when it is verified that some cognitive phenomenon only takes place, or does so with distinctive features, in conditions that are believed to favor (or inhibit) one type of processing or representation, conscious or unconscious, over the other. Say, certain properties of stimuli (e.g., distinctiveness, complexity, intensity, etc.) might be manipulated so that subjects taking part in an experiment will consistently claim not to perceive them; if despite this it can be shown that they are indeed capable of processing those stimuli, then it is believed that cognition dissociates at the level the stimuli are no longer conscious cognition; if a subject performs accurately in a task while her/his attention is directed elsewhere, then it is concluded that conscious and unconscious cognition are dissociable with respect to this cognitive factor.

2.2. Dissociation criteria, thresholds, measures, and tests

Falsifying the null model in an empirical framework requires that one be able to exhaust unconscious cognition (i.e., *all* unconscious cognition must be detected in a task) and to find conclusive evidence that is so exclusively for unconscious cognition (i.e., *only* unconscious cognition must be reflected in the output of some task). These are known in the literature as the *exhaustiveness* and *exclusiveness assumptions*, respectively (Reingold & Merikle, 1988; 1990); see also Schmidt & Vorberg, 2006). These assumptions motivate dissociation *criteria* and *thresholds*, as well as *measures* and *tests* of dissociation.

We begin with the dissociation criteria. A dissociation between conscious and unconscious cognition can be established according to the following general criteria:

- (a) Subjects are acquiring and/or applying information about stimuli or features thereof despite conceiving no conscious intention and/or no conscious strategy (control) to do so (the *intentionality criterion*; e.g., Graf & Schacter, 1985; Schacter, 1987; see also Jacoby, 1991; Jacoby et al., 1992);
- (b) Subjects are acquiring and/or applying knowledge of stimuli or features thereof without being aware of that fact (the *metaknowledge criterion*; e.g., Dienes & Berry, 1997; Dienes & Perner, 2002);
- (b.1) Subjects display an above-chance performance while claiming that they are merely guessing (the *guessing criterion*: introduced by Cheesman & Merikle, 1984; labeled by Dienes et al., 1995);
- (b.2) Subjects exhibit a mismatch between accuracy in the performance of a task and confidence in their own performance (the *zero-correlation criterion*: introduced by Chan, 1991; labeled by Dienes et al., 1995).
- (c) Subjects are incapable of reporting how and what they are acquiring/retrieving by means of free report or by any other verbal means (the *reportability criterion*; an all too diffuse criterion);

These criteria in turn require the establishing of *thresholds of awareness* (Cheesman & Merikle, 1984; 1986). Let *awareness* equate with the ability to make a discriminating verbal report:

A *subjective threshold* may be defined as the detection level at which subjects claim not to be able to discriminate perceptual information at better than a chance level, whereas an *objective threshold* is the detection level at which perceptual information is actually discriminated at a chance level. (Cheesman & Merikle, 1984, p. 391).

Cheesman and Merikle (1984, 1986) claim that it is on the subjective threshold that we must place the dissociation between conscious (above the subjective threshold) and unconscious (below the subjective threshold) discrimination of stimuli, and that therefore this is the level to take into consideration when experimenting into unconscious cognition. The objective threshold is the level *not* to take into consideration when trying to demonstrate unconscious cognition empirically (see Table 1).

These dissociating thresholds call for *dissociating measures* defined in relation with the two levels of awareness (see Table 1). *Subjective measures*, intended to measure the extent to which subjects process/represent more than they can report (introspect), rely on their reports concerning the extent of their (lack of) knowledge rather than on their reports on the discrimination of stimuli presented. In turn, *objective measures* are intended to measure how much people process/represent, and therefore ask the subjects to report on stimuli discrimination.

Yet another distinction, now having to do with testing, can be mapped into these dichotomies (see Table 1): *direct tests* ask the subjects to discriminate features of stimuli (e.g., grammatical vs. ungrammatical; left/right; color; etc.), whereas *indirect tests* make no reference to such features. In effect, direct tests rely on cognitive phenomena

such as conscious recognition (i.e., they directly test the cognitive phenomena: what/how much/etc. subjects consciously perceived, memorized, and can recall), while indirect tests probe the unreported effects of the cognitive phenomena by asking subjects for preferences and likings and by applying paradigms such as priming followed by, say, word-stem completion. For instance, if after being exposed to an audio list of words during anesthesia subjects tend to complete word stems with items from that list rather than with equally likely words, one can argue that they have been influenced/affected in the sense that those words became activated and were thus more readily used (see Reingold & Merikle, 1988, and Merikle & Reingold, 1991, for thorough discussions). This is the *qualitative difference* in stimulus effect that for some (e.g., Merikle & Daneman, 1998) allows us to conclude that there indeed is a dissociation between conscious and unconscious cognition.

Table 1: Dissociating awareness thresholds, measures, and tests. (*St* = stimulus)

		MEASURES		_
		Objective:	Subjective:	
		Discrimination	Introspection	
AWARENESS THRESHOLDS	Objective : subjects re- port that <i>St</i> was (not) presented	\checkmark		Direct: Discriminating responses
	Subjective: subjects can- not report that they perceived St		\checkmark	Indirect: Unre- ported or un- perceived ef- fects

TESTS

2.3. Cognitive formats/levels and factors

The *cognitive format* or *level* is key to the criteria, thresholds, measures and tests above: they focus both on the *processing* aspects of the cognitive phenomena of interest and on their *representation* features. The distinction between processes and representations is central in cognitive psychology and cognitive science (e.g., Gazzaniga et al., 2002; Thagard, 2005). Roughly, representations refer to mental contents, and processes to the operations or procedures that can be applied to them (e.g., Demetriou et al., 2010). Representations are, in other words, knowledge states or types (e.g., short- or long-term memory; explicit learning); knowledge is in the form of concepts, propositions, rules, images, etc., and learning, reasoning, and decision making are some of the cognitive actions or behaviors defined by the diverse procedures on them-most generally: acquisition, storage, and retrieval (Thagard, 2012; see Augusto, 2014, for a more formal approach). For instance, in order to make a decision on the shortest way to one's favorite coffee shop on campus, one needs to retrieve one's previously acquired and stored mental map (image) of the campus; additionally, one has to process perceptual information on one's present location, so as to update the mental map of the campus and one's position in it.

An important contribution of studies in unconscious cognition is the finding that the mental map of the campus (the representation), its retrieval from memory or one's perception of the present location (the processes), or even one's deciding on the shortest way (the overall cognitive behavior) need not be conscious: the representation can be implicit or procedural, the retrieval process can be purely automatic, one's present location can be unperceived (due to, for instance, inattention), and one may be wholly unaware of the decision making only to find oneself surprisingly in one's favorite coffee shop on campus.

Based on the distinction above between processes and representations, these various findings are accounted for by the processing and representational dissociations (see Table 2 in anticipation of the following elaboration). The processing dissociations are based on evidence that stimuli or properties thereof that are not consciously perceived by the subject trigger behavior that suggests that they are in fact processed, often against the subject's expectations or confidence. In accord with the exclusiveness and exhaustiveness assumptions, the stimuli are unperceived because they are subliminal (e.g., in priming) or unattended to (e.g., in classical conditioning), or due to some cognitive deficit of the subject (e.g., neural lesion) (see Sections 3.1-3). The representational dissociations, on the contrary, rely on evidence that specific features of perceived, attended to, stimuli cannot be identified or expressed (known) consciously, yet it can be verified that subjects do *represent* (know) them in an unconscious mode—i.e., they have been learned and memorized implicitly or procedurally (Sections 3.5-6 below). The dissociation covert vs. overt includes both aspects: covert behavior can be triggered by either unperceived stimuli, as in classical conditioning tasks, or perceived stimuli, as in prosopagnosia; in any case, the use of instrumentation is here required (Section 3.4).

Table 2: Dissociations between unconscious and conscious cognition according to cog

 nitive format or level.

COGNITIVE	CONSCIOUSNESS STATUS		
FORMAT / LEV-	Unconscious	Conscious	
EL			
	Independent; Dom-	Dependent; Subor-	
	inant	dinate	
Processing	Automatic	Controlled	
	Bottom-up	Top-down	
	Covert	Overt	
	Implicit	Explicit	
Representational	Procedural; Non-	Declarative	
	declarative		

In the representational paradigms, the stimuli must be perceived, because subjects must be allowed to form some positive knowledge of the stimuli—they know *what* the stimuli are—, though they are believed to fail to acquire conscious knowledge of some of the perceptual information. This entails that these paradigms raise issues that have mostly to do with the exclusiveness assumption. On the contrary, the processing paradigms are mostly challenged with respect to the exhaustiveness assumption. This means that the dissociation covert vs. overt motivates both issues.

This central distinction between representations and processes is not an irreconcilable foundational dichotomy.¹ In effect, these two formats or levels by and large share dissociating properties based on the *cognitive factors* seen as implicated in the cognitive phenomena of interest (see Table 3). It is largely upon these factors that the

¹ Contrarily to, for example, the distinctions process vs. product (e.g., Sun, 2008) and vehicle vs. process (e.g., Atkinson et al., 2000).

exhaustiveness and exclusiveness assumptions rest: they determine *all* the perceptual information in a stimulus that is unconscious, and they account for the perceptual information that is *only* unconsciously perceived or represented. For instance, at the representational level unconscious cognition is implicit because such properties as the high complexity of the stimuli (say, letter strings) and the need for a large storage space are believed to render the explicit format inadequate or simply not viable; these factors exhaust the unconscious properties of the stimulus in tasks of artificial grammars (see Section 3.5.1). These same properties are also often invoked to account for the bottom-up (vs. top-down) character of unconscious cognition (Section 3.3).

Table 3: Dissociations between unconscious and conscious cognition based on cognitive factors.

COGNITIVE	CONSCIOUSNESS STATUS		
FACTORS	Unconscious	Conscious	
Computation	Parallel; inexpen-	Serial; expensive;	
	sive; effortless	effortful	
Storage/Memory	Large; robust	Small; highly cor-	
		ruptible	
Complexity	High	Low	
Speed of processing	Fast	Slow	
Accuracy	Accurate	Inaccurate	
Flexibility	Inflexible; default	Flexible	
Attention	Inattentional	Attentional	
Capacity	Unlimited	Limited	
Executive control	Involuntary; unin- tended	Voluntary; intended	
Style	Divergent; rule-	Convergent; rule-	
	free; aschematic	constrained; sche-	
		matic	
Evolutionary status	Older; shared with	More recent; hu-	
	other animals	man-specific	
Context	Context-specific	Abstract	
Language	Nonverbal; non-	Verbal; proposi-	
	propositional	tional	

3. The main dissociation paradigms

3.1. Independent/dominant vs. dependent/subordinate processing

Unconscious cognitive processing is believed to be overall *dominant* and *independent* from the subjects' conscious processing of some task. By and large, dominance and/or independence of unconscious processing is invoked when it is seen that the cognitive processing of a certain input stimulus or features thereof (the actual performance) antagonizes, or belies, the subjects' confidence or conscious expectancy with respect to the same, i.e., the way they consciously process the task in cognitive terms. This paradigm is seminal in the scientific approach to unconscious cognition, having contributed to a long tradition of studies into "behavior without awareness" (Adams, 1957; Miller, 1939). Associative learning and some clinical conditions constitute today its main empirical settings, and the experimental apparatus typically relies in properties of stimuli that have to do with cognitive factors such as computation, accuracy, flexibility, and executive control (see Table 3).

Associative learning has long been seen as an adequate means to study unconscious cognitive processing (e.g., Lazarus & McCleary, 1951; McCleary & Lazarus, 1949). Classical conditioning is here especially important in that it is believed that subjects learn an expectation of the (non-)occurrence of an unconditioned stimulus (US; e.g., Reiss, 1980) that antagonizes the conscious learning of the same. For instance, Perruchet (1985) reported that subjects conditioned by a series of a conditioned stimulus (CS; a 70 dB sound) alone and in alternation with pairings (reinforcement) of CS and US (a puff of nitrogen delivered to the face) made increasingly conditioned responses

fitting sequential patterns in the presentation of US while exhibiting a decrease in expectation for it across runs of reinforced trials, i.e., they would blink even when they were consciously not expecting the US. This dissociation between the conscious expectation of an event and the dominant or independent unconscious processing of the same event is now referred to as the Perruchet effect (Perruchet, 1985; Perruchet et al., 2006; see also Perruchet, 2015), and replications of this effect, both with eye blinking as well as with other responses, have contributed significantly to the dual-process/system theory (e.g., McAndrews et al., 2012; Weidemann et al., 2009).

There is a plethora of clinical conditions in which there is a mismatch between actual performance and self-report regarding confidence or expectancy (see, e.g., Köhler & Moscovitch, 1997; Naccache, 2005). This is particularly manifest in research into blindsight (see, e.g., Weiskrantz, 1986; for a critical review: Cowey, 2010). In this condition, patients show remarkable visual capacities in what are called scotomata, or blind fields, which can in some cases occupy the left/right visual fields, or even the entire visual field. When asked to reach for and grab objects presented in their blind fields, or to navigate obstacles they claim not to see, patients often perform with remarkable accuracy (see, e.g., de Gelder et al., 2008), suggesting that unconscious processing of the stimuli is carried out despite the subject's conscious reluctance and/or lack of confidence.

A neural dissociation is often invoked to explain this dissociation: these patients are believed to have an impaired ventral stream, namely caused by lesions or malformations in the primary visual cortex (V1) that account for the more or less extensive scotomata; their good performance in object location and obstacle avoidance tasks is attributed to the preserved dorsal stream, which is believed to process visual stimuli independently of retino-cortical visual input and in an unconscious manner (e.g., Fang

& He, 2005; Milner, 2012; see also Hodges et al., 1999; Milner & Goodale, 2007). This is known as the dual visual stream hypothesis (DVSH). In fact, the processing in the dorsal stream appears to be dominant: for instance, healthy subjects are unable to inhibit corrections to perceived target perturbations and produce movement corrections to unperceived target perturbations (see Cameron et al, 2007). Also important, normal processing in the ventral stream appears to require focal attention (e.g., Treisman, 2006), being thus more dependent on attentional resources than the processing by the dorsal stream (see Table 3). Evolutionary aspects might be at play that not only make processing in the dorsal stream independent from the ventral stream, but perhaps dominant, namely due to evolutionarily older subcortical structures implicated in the dorsal stream (the superior colliculus and the pulvinar; see, e.g., Krauzlis et al., 2013; see Table 3).

Further neuropsychological evidence is available for this dissociation paradigm, as we now believe to have identified neural pathways and/or neurophysiological indicators of independent unconscious processing of masked words and in repetition priming (e.g., Dehaene, 2009; Dehaene et al., 2001).

3.1.1. Main issues

Despite this scientifically sound scenario, some words of caution are called for. To begin with, if more immediately taken within a framework opposing introspection and actual performance, then the well-known issues with the former are simply carried on into this paradigm; namely, the subject may confabulate, simply be wrong, or even hal-lucinate regarding their confidence and even their expectancies, i.e., regarding their higher-level mental states² (see, e.g., Dehaene & Naccache, 2001; but see also Dienes,

² The term 'higher-level' as applied to mental states is clarified below in Section 3.5.1.

2008, p. 53ff). As a matter of fact, it has more recently been claimed that confidence measures fail to probe/elicit the subjects' discrimination of properties (e.g., stimulus structure) that might better reveal the dissociation between conscious and unconscious cognition (see Dienes, 2012).

The belief that associative learning is wholly independent from the subject's conscious processing has been questioned by models that see conscious, language-based control (see Table 3) as necessary for this kind of learning to take place (e.g., Mitchell et al., 2009). In particular, it appears that conscious expectation is not so straightforwardly dependent or subordinate as the Perruchet effect claims it is, having actually a causal role in unconscious cognitive processing, namely by contributing to unconscious preparation for action (see Umbach et al., 2012).

Finally, techniques believed to give us access to (measures of) physiological indicators of (un)consciousness, in particular of the so-called neural correlates of consciousness, have been the focus of controversy, whether considered together with introspection or per se (see, e.g., Cleeremans & Haynes, 1999; Rees & Frith, 2007; Seth et al., 2010). In this context, the DVSH has been often challenged, with its main dissociations under severe critical scrutiny. For instance, Cardoso-Leite and Gorea (2010) claim that the conceptual fuzziness of such dissociations as consciousness/unconsciousness and perception/action, together with methodological issues that make it impossible to decide on whether the subjects are indeed hemianopic, render the DVSH basically untestable. Reviewing the existing evidence, they propose a single processing stream, dismissing thus the DVSH and its dissociating consequences.

Particularly afflicted by challenges more immediately respecting the exhaustiveness and exclusiveness assumptions is blindsight: is this phenomenon merely attributable to threshold-vision, spared islands of conscious vision, or even just an effect of scattered light? These (Campion et al., 1983) and further challenges (e.g., Fendrich et al., 1992) have not yet been definitely addressed.

3.2. Automatic vs. controlled processing

It is often the case that independent and dominant cognitive processes are automatic, but automaticity presents specific features; in effect, the present dissociation encompasses a larger domain of cognitive responses than the one above (see Section 3.1), with attention allocation and stimulus vs. context priority (see Table 3), for instance, as some of the main phenomena of interest.

Cognitive processes are said to be *automatic* when they are independent from attention but are context-dependent, they commonly cannot be stopped once they have been triggered, and they are believed to be inexpensive from the viewpoint of processing resources, having apparently unlimited capacity (e.g., Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Most importantly, they are believed to evade the subject's *control*. Cognitive processing believed to implicate control is seen as having features opposite to those above (see Table 3).

This paradigm became salient in cognitive psychology in the 1970s (e.g., Posner & Snyder, 1975) and is now firmly established in cognitive science (e.g., Schneider & Chein, 2003; Sun et al., 2009) and in social psychology, acquiring particular significance in the latter. Whereas in cognitive psychology automaticity and control are analyzed in the perception of stimuli, in social psychology researchers are typically more interested in the automatic/controlled influence that socially relevant stimuli perceived unconsciously have on social agents (e.g., Bargh & Williams, 2006). Because automa-

tism is here equated with unintentional and/or involuntary behavior (Bargh, 1994; Hassin et al., 2005; Uleman & Bargh, 1989), the activation of stereotypes and attitudes is believed to influence behavior more than when social constructs are processed in an intentional, controlled way (Fazio et al., 1986). The fact of the matter is that evaluations of social stimuli involving such theoretical constructs as stereotypes, attitudes, and selfesteem often implicate what appear to be wholly spontaneous responses that not only elude the conscious agency of the subject, but also are often demonstrably against the subjects' conscious explicit beliefs (e.g., Banaji & Hardin, 1996; Barden et al., 2004; Bargh et al., 1992; Fazio et al., 1986; Gilbert, 1989).

High on the list of dissociating tasks in this paradigm is priming. The central idea is that primes (words or pictures flashed briefly and immediately followed, or masked, by a target word/image) presented below 50 ms are often undetected by subjects in an experimental task, and are thus processed in an unconscious way. The typical experiment consists in priming subjects with specific social stereotypes (e.g., rudeness, old age, gender, race, etc.) and observing their subsequent behavior in situations where those stereotypes are believed to be automatically activated. For instance, Bargh and colleagues (Bargh et al., 1996) reported that in one of three experiments subjects who had been primed with the elderly stereotype spontaneously walked more slowly. Graham and Lowery (2004) reported that police officers in a racial priming condition (words related to the category *black*) showed overall more negative, stereotypical reactions to a hypothetical adolescent who had allegedly committed a crime. This is the automatic, i.e., unintended, passive and unconscious, influence of stereotype priming on behavior, according to these authors. These and similar effects elicited by priming in the context of the automaticity (vs. control) of social constructs, as well as of social evaluations (see Section 3.5.2 below), have been abundantly reported (e.g., Aarts et al., 2005;

Abbate et al., 2013; Blair & Banaji, 1996; Spruyt & Tibboel, 2015; Steele & Ambadi, 2006).

In a reading task, Stroop (1935) famously verified that the controlled process of naming ink colors is greatly hindered by the non-matching printed color names (for instance, the word *red* printed in blue ink) whose reading is automatically activated. Tasks using variations on the Stroop effect have also been devised to test automaticity in social constructs. In an auditory Stroop task, Most and colleagues (Most et al., 2007) reported that both adults and children were slower to categorize the sex of voices when there was sex-voice incongruence (e.g., a male voice saying the word *lipstick* or the name *Cindy*). The classical Stroop effect interpretation is here that the sex of the speaker and related stereotypes and attitudes are automatically activated in the listener, interfering with consciously controlled responses. Similar results are reported with racial stereotypes and prejudice (e.g., Richeson & Shelton, 2003).

Although these tasks use perceived stimuli, experiments combining the Stroop effect with priming corroborate their results. In evaluative semantic priming (e.g., Bargh et al., 1992; Fazio et al., 1986), responses to a target word were found to be facilitated in prime-target evaluative congruence. For instance, Greenwald and colleagues (Greenwald et al., 1989) presented briefly to the non-dominant eye of the subjects an evaluatively good-bad polarized priming (e.g., *fame*, *stress*) which was either backward-ly or simultaneously dichoptically masked by a random letter fragment pattern presented to the dominant eye. Their main task consisted in the subjects' deciding whether a series of target words was good or bad in terms of meaning. These authors verified that congruent trials (positive prime / positive target) elicited faster evaluative decisions in comparison with incongruent trials (e.g., negative prime followed by positive target). They grounded their results on the well-established fact that masked stimuli cannot be

detected, being thus unlikely to fall within the subjects' attentional focus; according to them, this detectionless processing paradigm provides a more adequate paradigm for probing unconscious cognition as compared to the mere diversion of the subjects' attention to another task (attentionless processing).

3.2.1. Main issues

The question is whether this can be described as automatic behavior in the senses above. Firstly, there seems to be a kinship between automatism and reflex behavior in humans, which puts the former at the level of low, basically sensorimotor, processes; on the contrary, social evaluations often are called for at the higher levels of cognition (e.g., decision making). Secondly, information processing operations can be said to be automatic in the sense above in particular when they are well- or over-learned. For instance, the activation of words in a semantic network or memory appears to be automatic in this sense: it is a very fast operation that commonly does not require attention or other cognitive expenditures. The decoding of phonological aspects of speech is even more telling of automaticity, as the processes by means of which the listener perceives and analyzes the sounds s/he hears are carried out without any apparent effort or control and are typically involuntary.

This opposition between automatic and controlled processes is made quite salient in the Stroop effect, but this and the above are all processes that are greatly invariable from subject to subject, suggesting that, beyond learning and practice, there is also some innateness involved (see, e.g., Hasher & Zacks, 1979). On the contrary, the activation of attitudes and stereotypes, though clearly fast and effortless, shows too large an inter-subject variability to be seen as merely automatic (e.g., Devine & Sharp, 2009).

In effect, what an abundance of studies has reported (see, e.g., Greenwald & Banaji, 1995; Payne & Gawronski, 2010) is that these higher thought processes are often unconscious, but this does not allow us to speak straightforwardly of automatisms (e.g., Dijksterhuis, 2010). Even when automaticity does appear to be allied with unconscious processing, it is so in more than one way or degree (Bargh, 1989; 1994). Furthermore, automaticity is not an all-or-none feature with respect to psychological phenomena, allowing partial control (e.g., Logan, 1989). In turn, control may be exercised unconsciously (e.g., Moskowitz, 2001; Moskowitz et al., 1999). These results in social psychology are largely corroborated by recent findings in cognitive psychology (e.g., Dehaene et al., 2006; Martens & Kiefer, 2009; Martens et al., 2011; Kiefer, 2007, 2012); in particular, cognitive factors such as attention are now believed to be implicated in both conscious and unconscious processing modes, albeit in different ways (e.g., VanRullen & Koch, 2005). Finally, highly automatized and often unconscious actions such as typing and driving a car belie the equation between automaticity and lack of intention or will: one typically engages in and performs such actions intentionally (i.e., with a goal) and willingly.

The consequences of this from the perspective of social cognition are vast, impacting on our perception of social phenomena such as stereotyping, prejudice, and discrimination (see, e.g., Fiske, 1989; see Devine & Sharp, 2009, for a review). These implicate the social agent in a crucial way, and the purported unconstrained character of the activation of stereotypes and attitudes clashes with our notions of agency, free will, and social responsibility. More specifically, seeing these social phenomena from a rigid dual-process/system perspective might lead to theories of unintended and/or involuntary thinking in social psychology; this might contribute to the view that social agents are inevitably incapable of controlling their social constructs and the unconditional auto-

matic activation thereof (Bargh, 1999). However, recent work has reported failure to replicate some of the results above, as well as in cognitive psychology. For instance, Doyen and colleagues (Doyen et al., 2012) were unable to replicate the elderly priming effect of Bargh and colleagues (Bargh et al., 1996), and another study (Harris et al., 2013) reported failure to replicate findings of automatic activation of high-performance goals also in the priming paradigm (see Bargh et al., 2001). Although this has generated an ongoing debate (e.g., Cesario, 2014), it nevertheless casts additional doubt on the automaticity vs. control dissociation in cognitive processing.

In fact, a thorough revision of the theory and concepts in the research into automaticity appears to be now a priority (e.g., Evans & Stanovich, 2013; Moors & De Houwer, 2006).

3.3. Bottom-up vs. top-down processing

The distinction between bottom-up and top-down processes is central in psychology and cognitive science (e.g., Coon & Mitterer, 2010, p. 151ff). Roughly, the former processes are stimuli-driven, whereas the latter are knowledge and expectation-based. In the context of unconscious cognition, bottom-up processes characterize unconscious processing and top-down processes are claimed to do so for conscious cognition (see, e.g., Sun, 1999). This distinction is highly diffused in many dissociative approaches (see Sun et al., 2001), but in the case of unconscious thought theory (UTT), it plays a role so fundamental as to be a principle: the bottom-up-versus-top-down principle (see Bos & Dijksterhuis, 2011; Dijksterhuis & Nordgren, 2006).

UTT's theoretical ground and methodological approach can be encapsulated in a few principles (Dijksterhuis & Nordgren, 2006). Divert one's attention, leave it to the

unconscious to make the complex decisions one is faced with (i.e., let it "incubate"), and one is much better off, because the unconscious has a much higher processing capacity as compared to conscious thought (the capacity principle). Although it cannot follow rules, it delivers an estimate weighted (the weighting principle) and rough but accurate (the rule principle). This processing is characterized by a divergent search that is more extensive and encompassing than the one the convergent-like conscious thinking can offer (the convergence-versus-divergence principle).

It is not clear whether these principles follow from, or support, another, most elemental principle, to wit, the Unconscious Thought principle (authors' capitals), according to which there are two kinds of thought different in nature, viz., conscious and unconscious thought. These principles are motivated by the deliberation-without-attention effect, stated as the hypothesis that better effects ensue from non-attentive (i.e., unconscious) deliberation. The whole theory is crowned by UTT's *pièce de résistance*, the already mentioned bottom-up-versus-top-down principle, according to which unconscious thinking is aschematic, whereas conscious thought works schematically. It is claimed that by avoiding schemas and like hindrances, unconscious thinking is capable of integrating information in a better way, as well as of avoiding biases that characterize conscious thought (Bos & Dijksterhuis, 2011).

3.3.1. Main issues

Scientifically, UTT might be a rather naïve theory, a potpourri of theoretical principles with very little flesh to their bones according to some, and all too easily falsifiable according to others (e.g., Acker, 2008; González-Vallejo et al., 2008; Huizenga et al., 2012; Srinivasan & Mukherjee, 2010; see Bargh, 2011, for a discordant voice). But in

spite of this criticism, which also draws attention to the fact that its reported results often fail to be replicated, UTT highlights what in many other approaches has remained an unspoken assumption, to wit, that unconscious processing is exclusively or overwhelmingly default or bottom-up (i.e., basically stimulus-driven and aschematic) whereas its counterpart is top-down, schematic and/or context-driven (*but* see Table 3). However, just as there are apparently substantiated claims that unconscious processing can be partly or even greatly top-down modulated or influenced (see above), there are also like claims that context-driven, top-down processes can be partly, greatly, or even wholly, unconscious (e.g., Van Opstal et al., 2011). In particular, it appears that both processes interact in attentional performance (Egeth & Yantis, 1997). In other words, it has been partly accepted that unconscious processing can go as "up" as to involve, say, semantic processing, and it can be as top-down-like as to control cognitive processes, and this dissociation is fast losing the little ground it might have gained.

This loss is at the same time a result of, and entails the loss of favor of, other related dissociations, such as the dissociation between implicit and explicit cognitive formats or levels, to be approached below in Section 3.5 (see, e.g., Sun & Zhang, 2004). This is accounted for by the fact that UTT actually sins in presenting subjects with perceivable stimuli, claiming that they are however unattended. This means that in fact subjects are capable of *explicitly* representing the stimuli, but some properties thereof are believed to remain *implicit* due to inattention. In other words, this is the implicit vs. explicit dissociation, and the UTT's appeal to a bottom-up vs. top-down processing becomes superfluous, if not altogether vacuous. In effect, by using perceivable stimuli any task in this paradigm might be considered as one in bottom-up vs. top-down *conscious* processing modes (see introductory paragraph to this Section).

3.4. Covert vs. overt cognitive processing/representations

We speak of *overt* behavior (e.g., verbal report) for indications of (non-)discrimination, (non-)recognition, (non-)preference, (non-)familiarity, etc. By *covert* we here mean that a cognitive phenomenon is inaccessible to both private and public observation without the use of instrumentation. The rationale for this dissociation paradigm is as follows: if the subject does not recognize/identify/etc., a stimulus overtly and yet it can be shown by instrumental means that they react physiologically to the same more or less vigorously, then it is believed that this covert behavior expresses unconscious cognition of the stimulus in question.

This dissociation is based on our present understanding of the neural bases of human information processing, in which emotion plays a central role (e.g., Tamietto & de Gelder, 2010). The range of physiological measures for covert behavior is quite large, encompassing from electrophysiological to hemodynamic components of neural events (see Andreassi, 2007). Examples of the former measures are the electroencephalogram (EEG) and its associated event-related potentials (ERPs), and functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) are examples of the latter. A well-established measure of covert behavior in this dissociation paradigm is the skin conductance response (SCR), a measure of electrodermal activity at the level of the autonomic nervous system believed to be associated with emotional valence of stimuli and tasks.

The visual agnosias are paramount in this paradigm (see Augusto, 2010) and SCRs have long been seen as an appropriate dissociation measure (e.g., Bauer, 1984). A particularly interesting visual agnosia, given the exceptional importance of faces in the study of unconscious cognition (Axelrod et al., 2015), is prosopagnosia. This is a condi-

tion characterized by the incapacity of the individuals afflicted to recognize faces, including those of very familiar people, and often even their own, despite exhibiting no degradation in visual abilities or in other cognitive skills. Nevertheless, today evidence abounds, including on double dissociations, that these individuals produce significant SCRs when presented photos of the faces of people who are close to them (e.g., Damásio et al., 1990; Tranel et al., 1995). One of these double dissociations is particularly interesting in that it suggests a "mirror image" condition of prosopagnosia: in the Capgras delusion, it is precisely the autonomic, covert recognition of people that is impaired (SCRs are absent), leaving the patients with the distressing experience of consciously recognizing persons close to them and yet claiming that they are impostors (e.g., Ellis & Young, 1990; Ellis et al., 1997). It is possible that these two conditions reflect a double dissociation that can be mapped into a limbic dual visual system, a hypothesis first conceived by Bauer (1984): in the one case, prosopagnosia, only the neuroanatomical pathway of covert processing of emotional content, to wit, the "dorsal visual-limbic pathway" (DVLP), is unimpaired, with deficit occurring in the "ventral visual-limbic pathway" (VVLP); in the mirror case, the Capgras delusion, the unconscious, covert emotional processing is believed to be greatly or entirely compromised due to damage to the DVLP, whereas the overt recognition in the VVLP is preserved (see Ellis & Lewis, 2001, for neuroanatomical details).

This dissociation is also explored by means of SCRs at the executive level. The somatic marker hypothesis (SMH) explores the dissociation covert vs. overt in decision making. The SMH posits the ventromedial prefrontal cortex (VMPFC) as the locus of covert, greatly or wholly unconscious learning responses to good and bad outcomes of one's decisions when uncertainty is involved (Bechara & Damásio, 2005; Damásio, 1994; Damásio et al., 1991). In the associated Iowa Gambling Task (IGT), it has been

consistently reported that normals produce substantially higher anticipatory SCRs when preparing to make a "bad move," accounting for beneficial gambling behavior even when the gambler is completely unaware with regard to what is actually going on in the game—what appears to occur with 30% of the participants. On the contrary, patients with lesions to the VMPFC fail to produce SCRs and perform poorly in the IGT, consistently making bad decisions (see, e.g., Bechara et al., 1997). Damásio and colleagues conclude that the somatic markers that covertly "mark" the cognitive processing, i.e., neurotransmitters released by the central nervous system, are not activated (*minded*, as they put it), and thus fail to covertly "inform" the patients via anticipatory SCRs that their decisions are prejudicial in the context of the IGT.

3.4.1. Main issues

Recruiting a heavy neurophysiological apparatus, this paradigm has the advantage of approaching unconscious cognition in a very objective way. This, however, is not problem-free, with the meaning and usage of covert measures motivating conflicting issues, many of which with foundational implications (see, e.g., Ortu, 2012). In effect, it is often suspected that correlations between neurophysiological measures and cognitive phenomena might be misleading (e.g., SCRs appear to be correlated with perhaps too many and too diverse emotional responses; see, e.g., Figner & Murphy, 2011), and measuring devices may be inaccurate or simply inadequate, generally or contextually so. Also, these measurements are indirect and often peripheral: for instance, fMRI measures brain activity by indirectly reading the brain's hemodynamics, and SCRs are measures of electrical conductance in the skin, a peripheral phenomenon of sweat production. (See Brouwer et al., 2015, for these and further issues.)

The above issues pose major concerns, given that the use of instrumentation is essential in segregating this dissociation paradigm, but the distinction between covert and overt behavior in the general sense is also not always clear-cut. For instance, intestinal movements, commonly covert, can become audible (borborygmi) and thus overt without for that entailing any essential changes (say, in intensity) that might have psychological significance. Even more tellingly, SCRs just are instrumental readings of autonomic sympathetic changes that alter sweat production, which can be actually overt in some environments, contexts, and/or people. All these and further issues plague the SMH, making it a highly controversial hypothesis (see, e.g., Dunn et al., 2006).

3.5. Implicit vs. explicit representations

The dissociation implicit vs. explicit was first devised in the early 1920s by McDougall (1924) as a distinction of both *knowledge* and *memory*, as well as of purpose and motivation. It was famously re-inaugurated in the late 1960s by A. Reber in the field of *learning* (Reber, 1967). In the middle 1980s, Graf and Schacter (1985) introduced the distinction into *memory* studies, using the label *implicit memory* for the type of memory that is tapped by priming tests, and referring to the kind of memory that is tapped by tasks of recall and recognition as *explicit memory* (see Schacter, 1987, for a history of both kinds of tasks). The essentially representational character of this dissociation (see Section 2.3) has thus been a constant. In the 1990s, this dissociation was taken by social psychology, where it acquired some distinctive features and raised new, specific issues. We thus discuss this paradigm in cognitive psychology and social psychology in two separate subsections.

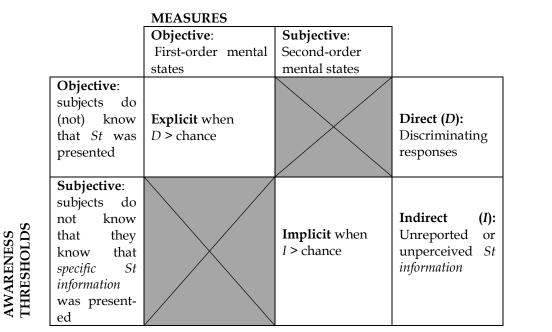
3.5.1. Implicitness vs. explicitness in cognitive psychology

Broadly, implicitness in learning and memory is said to be verified when people learn about—often complex—stimuli features and apply that knowledge without having the conscious intention of doing so and/or without being capable of consciously representing the knowledge acquired (cf. Dienes & Berry, 1997). These findings have been reported in tasks using sequence learning, artificial grammars, and complex simulated systems in the context of learning, and, now in the context of memory, in tasks of priming and skill performance (for reviews, see, e.g., Augusto, 2010; Cleeremans, 1997; Schacter, 1992; Shanks, 2005).

Because this dissociation focuses cognitive representation, on the metaknowledge criterion is here prominent: what is explored is basically a higher-order mental state of ignorance with regard to a state of positive knowledge in a lower-order mental state. The guessing and the zero-correlation criteria are of particular interest (see Section 2.2). In the guessing criterion, there is a lack of correspondence between the subjects' first-order mental state of confidence and their second-order mental state that they are merely guessing, whereas in the zero-correlation criterion a mismatch is verified between a first-order mental state of knowing (the subject makes the right responses) and a second-order mental state of ignorance. The rationale, one can argue, is that the more perceptual evidence one has, the more accurate one's responses will be, and the more this perceptual evidence is conscious, the greater one's confidence will be; high accuracy in face of lack of confidence implies that informational content of the stimuli remains unconscious for the subject (Dienes, 2008; Dienes & Perner, 2004; Ziori & Dienes, 2008).

In accord with these criteria, the subjective awareness threshold is deemed appropriate to approach learning and memory of the implicit kind, and therefore subjective measures and indirect tests are here prescribed: tasks should probe information that the subjects do not know that they know by testing for stimuli effects that are unreported or appear to go unperceived by the subjects, i.e. by probing second-order mental states of ignorance or lack of confidence with respect to first-order mental states. Subjects are said to have implicit representations when it is quantitatively verified that the results in the indirect tests are greater than chance and/or when the guessing and the zerocorrelation criteria are satisfied (see Table 4; compare with Table 1).

Table 4: Representational format implicit vs. explicit with relation to awareness thresholds, measures, and tests. (St = stimulus)



TESTS

Interestingly enough, the very first empirical study into unconscious cognition established this framework in large measure: Peirce and Jastrow (1884) found that differences in weight so minimal that the subjects (the authors themselves) simply could not discriminate consciously were in fact represented in an unconscious mode. The authors argued that if they did not actually know the difference, then their performance should be at chance level (50% correct answers); however, their performance was well above chance, while they could not report a difference in stimulus presentation. The bases were thus established for further, posterior work in the criteria, thresholds, measures, and tests for dissociation.

Recent support for this dissociation comes in the form of neurophysiological evidence for segregated neural paths for unconscious learning and retrieval of complex rules (e.g., Reber & Squire, 1994; Seger et al., 2000; Skosnik et al., 2002).

3.5.1.1. Main issues

It is only fair to say that much of the work into the criteria, thresholds, measures, and tests of dissociation was done in direct association with this paradigm (Section 2.2). However, some important issues remain.

Because stimuli are here perceived by the subjects, the exclusiveness assumption motivates a number of issues. One of the cruces in implicit learning and memory lies in the difficulty to determine whether the learned/memorized information was so unconsciously, or whether it just happens that subjects cannot consciously or intentionally retrieve at a later time what was in fact a conscious learning/memorizing episode (e.g., Destrebecqz & Peigneux, 2005; Destrebecqz et al., 2015; Dulany et al., 1984).

Further issues to do with the measuring and testing of implicitness in cognition contribute to the scenario of controversy and instability that surround this dissociation. For instance, the reliability of measures and tests of implicit cognition, namely of implicit memory, has been questioned, not the least because they often are less stringent compared with measures of explicit cognition (see, e.g., Buchner & Wippich, 2000). This is related to the problem of differential test sensitivity that comes up mostly with objective thresholds and the zero-correlation criterion, but which can be a general issue, as discussed by Shanks and St. John (1994) (see Dienes, 2011, for a proposed solution).

Although there has been an effort to establish criteria of implicitness (see Section 2.2 above), no such effort has been made with respect to explicitness in cognition. In fact, if we see this dissociation as one mostly between first- and second-order mental states, then explicitness might simply mean that subjects have veridical second-order mental states about their first-order mental states. The apparently intractable subject of introspection shows that such a property of mental representations is at best problematic, not the least due to biases and interpretative issues (Eriksen, 1960; Holender, 1986; Merikle, 1992). On the other hand, to define explicitness according to Table 4 is simply uninformative as far as unconscious cognition is concerned, namely because it remains at the level of lower-order representations.

The nail in the coffin of this dissociation threatens, however, to be put by the many terminological issues that plague it (e.g., Reingold & Merikle, 1990). The objective should be to avoid or entirely escape terminological issues by addressing directly the criteria underlying their usage. In particular, in our view the use of the term 'implicit' to refer to results obtained in tasks in which the stimuli are unperceived, such as in priming, contribute significantly to the terminological and conceptual chaos. In effect, in this dissociation paradigm we check for *specific unreported* or *unperceived* infor-

mation with regard to stimuli that the subjects *know were presented*; this is the indirect testing, at the subjective or second-order mental level, of the knowledge state formed below the subjective awareness threshold (see Table 4). In other words, to expect subjects to form a higher-order mental state of knowledge with respect to a stimulus that is unperceived—i.e., the subjects do not know what the stimulus is—poses problems now with regard to the exhaustiveness assumption (see e.g., Destrebecqz et al., 2015).

3.5.2. Implicitness vs. explicitness in social psychology

It is no wonder that social psychology, again, is, together with cognitive psychology, a field that invests greatly in representational implicitness: people seem to judge other individuals—and themselves—in many and disparate ways (from physical features, namely facial ones; from gestures; etc.) while being unaware both that they are judging others/themselves and that they are drawing on some knowledge base. It is now almost folklore that the acquisition of new biases and the formation of new attitudes and stereo-types can be exceedingly fast, and the activation of stored ones requires but little stimulation, with priming being sufficient to trigger social judgments and decisions. With disfavor falling on automaticity as an explanation for these phenomena (see Section 3.2 above), the notion of implicitness in cognition greatly appealed to social psychologists.

The change of focus in unconscious social cognition from automatic behavior to implicit cognition can be attributed to Greenwald and Banaji (1995). In this paper, the authors surveyed the main constructs of social psychology, to wit, attitudes, stereotypes, and self-esteem, from the unifying viewpoint of *implicit construct*, "the introspectively unidentified (or inaccurately identified) trace of past experience that mediates R" (Greenwald & Banaji, 1995, p. 5), where R names categories of responses. They accord-

ingly revised such important findings in social psychology as the halo effect, the mere exposure effect, the minimal group effect, etc., in light of the theory of implicit cognition, claiming that implicitness of the cognitive processes involved can be detected in two ways: firstly, when subjects attribute their responses to the wrong stimuli or to the wrong properties thereof, and, secondly, when the effects weaken or disappear when subjects are told what the implicit constructs and/or stimuli at play are.

Refinement was needed, and this was claimed to have been accomplished with the Implicit Association Test (IAT; Greenwald et al., 1998). Basically, the IAT is an association latency task: given two concepts (e.g., flower names vs. insect names) and an attribute (e.g., pleasant vs. unpleasant words), it is predicted that the association flower + pleasant should be easier (and therefore faster) than the association insect + *pleasant*, because the first is common, whereas the second is rather atypical. (Of course, just the reverse might be the case in particular cultural contexts.) Greenwald and colleagues assumed that a similar difference in performance in tasks which associated, say, white faces vs. black faces and an attribute, reliably and validly measured the strength of implicit associations between race and attributes. The authors of the IAT actually spoke of "associative structures" (*ibid.*, p. 1464) and cultural "immersion" (p. 1470), suggesting that implicit attitudes are socially- or culturally-caused components of personality that escape control by the individuals (see Section 3.2 above). In effect, the IAT is claimed to reveal the subjects' implicit attitudes whether or not they prefer to express them. Also importantly,³ the IAT is also claimed to measure implicit stereotypes and self-concepts by making some minor adaptations. The IAT was seen by its authors as a better method of investigating implicit attitudes than other methods, namely evaluative semantic priming (see Section 3.2 above).

³ Although stated in a footnote (*ibid.*, p. 1466).

A more recent task is the Go/No-Go Association Task (GNAT) (Nosek & Banaji, 2001). Just like the IAT and other priming tasks (e.g., Draine & Greenwald, 1998), the GNAT assesses the strength of association between a target category and a polarized attribute dimension. This paradigm claims to trade off response latency by sensitivity by evaluating the degree to which items belonging to the target category and a attribute can be discriminated from items not belonging to these (distracters). For instance, *fruit* is a positive category target; accuracy in discriminating fruit and good items from distracters should be higher than accuracy in discriminating fruit and bad items from distracters. This difference between the two conditions is taken to measure implicit attitudes or beliefs. Sensitivity is actually measured by how the subjects perform in their response tasks to briefly presented stimuli: "Go" (press a bar) vs. "No-Go" (do not press any key), when the presented target category and attribute match (e.g., *fruit* + *good*) and when items appear that do not belong to the target category and attribute (noise; e.g., *fruit* + *bad*), respectively. The socially significant experiments use, for example, race or gender category targets.

3.5.2.1. Main issues

The tasks above that are claimed to measure implicit social constructs fail not only to give effective definitions of implicitness, but also altogether omit definitions of explicitness (for instance, *explicit* may just be a synonym for *self-report*). By and large, implicit representations (i.e., attitudes, stereotypes, and self-concepts) are merely equated with unidentified, undetected, or inadequately identified/detected representations, and the matter is left at that. This accounts for a confusion between implicitness and automaticity: the IAT is in fact a return to the automaticity vs. control dissociation, in that

implicit associations are seen as being (strong enough so as to be) automatically triggered. This justifies the use of a latency task in the IAT. The same can be said of the GNAT, with the aggravating factor that the stimuli are only briefly presented.

The few attempts to define implicitness in this dissociation paradigm are unsatisfactory. As stated by Greenwald and Banaji (1995), implicitness in social cognition stands for the influence of past experiences in the behavior of a subject unaware of such an influence. This is to say that mostly implicit memory is addressed in this social paradigm, with little or no reference to implicit learning. In fact, often no account is given of how the constructs recalled in an implicit way might have been themselves implicitly learned. The consequence of this is clear: the model of implicit social cognition, and in particular the IAT, fails to distinguish what in the subject's behavior is purely and simply learned (an implicit social knowledge base) and what is actually believed (the subject's own unconscious attitudes and stereotypes). Moreover, from the latter, the IAT cannot distinguish those that are due to mere familiarity or exposure from those that are actively nourished by the subjects. (Nosek et al., 2007, (self-)review the IAT; for a comprehensive hetero-review and criticism, see, e.g., Fiedler et al., 2006.)

In sum, this dissociation paradigm operates with minimal and/or often inadequate definitions of the dissociation implicit vs. explicit, making it rather uninformative (e.g., Reder et al., 2009; Roediger, 2003; Willingham & Preuss, 1995). Furthermore, tasks in this paradigm have been challenged with regard to issues such as validity and reliability (e.g., Brendl et al., 2001; Fazio & Olson, 2003; see Perugini et al., 2010, for an in-depth study; but see Gawronski, 2009, for a defense and further paradigms). Moreover, the extent to which the dissociation implicit vs. explicit might be differential in social terms is not approached, that we know of (see also *Main issues* of Section 3.2).

3.6. Procedural/non-declarative vs. declarative representations

Perceived to some extent as secondary with relation to the dissociation implicit vs. explicit, this dissociation is in fact distinct: whereas the former dissociates between mental states (first- vs. second-order), this focuses on the modes of expression (language vs. motor skill). It has acquired an important status in cognitive science, where it features in some influential cognitive architectures such as Soar (e.g., Newell, 1990), ACT-R (e.g., Anderson, 1983; 1993), and CLARION (e.g., Sun, 2002). It is intimately associated with the multiple memory system (MMS) model that distinguishes two main components to long-term memory, declarative and non-declarative memory (see, e.g., Milner et al., 1998; Squire & Wixted, 2011).

As the terms 'declarative' and 'non-declarative'/ 'procedural' indicate, it is believed that the former cognitive format contrasts with the latter in that this cannot be expressed in verbal terms, being "executable" rather than verbalizable. In other words, we have here a dissociation between propositional representations and motor skills. To put it in simplistic terms, this is a dissociation between "knowing that" and "knowing how" (Cohen & Squire, 1980).

Riding a bike is often given as an example of *non-declarative/procedural memory*: if one asks any skilled biker how s/he rides a bike, the biker will be at a loss for words as to how s/he represents this skill, though s/he will readily show how to do it. The fact that s/he does it skillfully implies a learning period in which many actions were newly learned or adapted, and were practiced to the level of automatisms; some of these actions are indeed hard or impossible to express verbally (for instance, how one keeps one's balance on a bike), but they can nevertheless be more or less referred to otherwise, teaching this skill would be impossible. Contrastingly, *declarative memory* is

believed to be composed of facts (semantic memory) and events (episodic memory). These need not be separate, as one can know a fact and at the same time be capable of recalling the event in which it was learned (where, when, etc.), but clinical studies show that they can indeed be independent (e.g., Schmolck et al., 2002). The main aspect that unites these two kinds of memory is that their contents appear to be optimally retrieved and expressed in propositions (e.g., "Apples are a kind of fruit; "I first met him at the lobby of the hotel in 1980").

Behavioral and neurological evidence for a dissociation between the two major kinds of memory abounds, with emphasis placed on the fact that amnesics commonly seem to preserve features of non-declarative memory, while being impaired in tasks of declarative memory. In fact, amnesics not only appear not to lose contents from their non-declarative memory subsystems, but can actually learn new skills (Tranel et al., 1994) and associations (namely by conditioning; see, e.g., Daum et al., 1989; Weiskrantz & Warrington, 1979), and they even show unimpaired performance in priming tasks when indirect measures are used (e.g., Graf et al., 1984).

3.6.1. Main issues

The problems begin right here, however, since normal performance in priming tasks likely entails storage of the declarative kind, namely semantic memory: if priming is verified, then the subject must have acquired and stored information about, say, a list of words (e.g., Tulving et al., 1991), even if this does not involve learning new associations (e.g., Shimamura & Squire, 1989) and episodic memory of the event is poor or null.

This suggests that memory might be a more distributed and/or interactive phenomenon than this dissociation postulates. In effect, it is often the case that severe impairments in one of these two major memory systems caused by a plethora of conditions are accompanied by deficits in the other, even if these are mild or only occasional in occurrence (see, e.g., Budson, 2009, p. 72).

Nevertheless, the weight of evidence seems to point to a functional-anatomical separation between the two major memory systems postulated by the MMS model. But to paste this separation into the conscious vs. unconscious information processing dissociation is an altogether different business. To begin with, many forms of procedural or non-declarative memory are accessible to conscious recollection, though they might defy verbalization, even when they have been learned and stored unconsciously; this appears to be the case in several forms of conditioning (see Swift, 2003, p. 7). On the other hand, facts and life events that have been learned or experienced in a conscious state might become unconscious due to, say, trauma or feelings of guilt, and be recalled only implicitly, i.e., affecting behavior or in altered states of consciousness such as dreaming or hypnosis. Priming paradigms, in particular, threaten to blur the neatness of the MMS model in that, though apparently an implicit memory phenomenon, priming draws on semantic memory; the same can be said of subliminal perception. Moreover, recent functional-anatomical evidence offers a picture where the same brain loci/structures mediate both conscious and unconscious memory forms, and it has been known for long that the same memory functions can be highly distributed. In face of this picture, research paths (should) remain open, suggesting for some the dismissal of the consciousness vs. unconsciousness criterion in favor of other explanatory models of memory (e.g., Henke, 2010).

The procedural vs. declarative memory dichotomy was early on coupled to a distinction made between implicit and explicit memory systems (Graf & Schacter, 1985), so that today it is fairly common to see scientists interchanging the declarative vs. procedural and the explicit vs. implicit distinctions. However, just as the *consciousness status* of memory does not correspond harmoniously to a determinate *kind* — i.e., declarative memory is not necessarily conscious, and non-declarative/procedural memory is not necessarily unconscious —, so explicit memory need not be (both) declarative and conscious, and implicit memory need not be (both) nondeclarative/procedural and unconscious. In other words, there is no fixed correspondence between the *mode of expression* and the former two other aspects of memory (see also Augusto, 2013). In fact, all combinations are possible; for instance, a piece of declarative memory may be unconsciously elicited either in an implicit (e.g., in a preferences task) or in an explicit mode (e.g., in an artificial grammar task).

Finally, many of the findings reported in this dissociation issue from tasks in which stimuli are unperceived, as in priming, in conditioning, and in subliminal perception tasks (see last paragraph of the introductory section above). This poses major issues from the viewpoint of our methodological segregation based on the distinction in cognitive formats or levels (Section 2.3). In particular, there is the risk of confusing a procedural representation with the fact that the subjects simply had no sufficient perception of the stimulus to be able to form some positive knowledge of the same, i.e., they did not go beyond the processing level or format with respect to the unperceived stimulus.

4. Concluding remarks

The obvious way to approach unconscious cognition seems to be by how it dissociates from conscious processes and representations. Nevertheless, it has been historically an impressively difficult task to establish criteria, measures, thresholds, and tests for unconscious cognition based on this dissociation logic. Despite these, many issues remain, not the least the terminological and conceptual chaos that plagues the literature. In particular, there are still too many dissociations around, and often the claimed oppositions appear not be veridical under closer scrutiny. All this leaves those interested in the field of unconscious cognition lost in dissociation.

This paper is expected to contribute to alter this situation by addressing some of the issues above. We identify six main dissociations and distinguish them according to two cognitive formats or levels, to wit, processing and representational. Although this is a central distinction in psychology and cognitive science, it has not been explored in the field of unconscious cognition; we propose it be central in this field, too. One of the advantages is that methods are now segregated according to these two formats or levels; namely, unperceived stimuli (e.g., in priming, in classical conditioning) should be used only in the processing dissociations, whereas the representational dissociations must present perceivable stimuli. This helps to reduce and better tackle the challenges motivated by the fundamental assumptions in unconscious cognition, to wit, the exhaustiveness and exclusiveness assumptions, and eliminates the promiscuity of methods that accounts for much of the confusion in the field. Moreover, we significantly reduce the number of actual dissociations in the literature, contributing to a clearer view of both the trees and the forest.

We also make a distinction of secondary dissociations with relation to cognitive factors, and subordinate these to the two mentioned cognitive formats or levels. We further proceed to analyze critically the remaining main dissociations, and conclude that in fact they are not as clearly and rigorously established as required in a scientific field. This is not meant to dismiss the main dissociations *tout court*. On the contrary, the identification of the main issues that afflict them aims to motivate a more rigorous theoretical and methodological framework, and thus contribute to further, more solid research into unconscious cognition.

Acknowledgments: This paper was partly funded by a postdoctoral fellowship granted by FCT. Part of the research for this work was carried out at the University of Barcelona, where the author was a visiting researcher in 2015.

References

Aarts, H., Chartrand, T. L., Custers, R., Danner, U., Dik, G., Jefferis, V. E. & Cheng, C.L. (2005). Social stereotypes and automatic goal pursuit. *Social Cognition*, 23, 465-490.

Abbate, C. S., Ruggieri, S., & Boca, S. (2013). Automatic influences of priming on prosocial behavior. *Europe's Journal of Psychology*, 9. doi:10.5964/ejop.v9i3.603

Acker, F. (2008). New findings on unconscious versus conscious thought in decision making: Additional empirical data and meta-analysis. *Judgment and Decision Making*, *3*, 292-303.

Adams, J. K. (1957). Laboratory studies of behavior without awareness. *Psychological Bulletin*, *54*, 383-405.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). Rules of the mind. Hillsdale, NJ: Erlbaum Associates.

Andreassi, J. L. (2007). *Psychophysiology: Human behavior and physiological response*. Mahwah, NJ: Lawrence Erlbaum Associates.

Atkinson, A., Thomas, M., & Cleeremans, A. (2000). Consciousness: Mapping the theoretical landscape. *Trends in Cognitive Science*, *4*, 372-382.

Augusto, L. M. (2010). Unconscious cognition: A survey. Advances in Cognitive Psychology, 6, 116-141.

Augusto, L. M. (2013). Unconscious representations 1: Belying the traditional model of human cognition. *Axiomathes*, *23*, 645-663.

Augusto, L. M. (2014). Unconscious representations 2: Towards an integrated cognitive architecture. *Axiomathes*, *24*, 19-23.

Axelrod, V., Bar, M., & Rees, G. (2015). Exploring the unconscious using faces. *Trends in Cognitive Sciences*, *19*, 35-45.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, 7, 136-141.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, 87, 5-22.

Bargh, J. A. (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 / London: The 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, NJ: Lawrence Erlbaum Associates.

Bargh, J. A. (1999). The cognitive monster: The case against controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (eds.), *Dual-process theories in social psychology* (pp. 361-382). New York & London: The Guilford Press.

Bargh, J. A. (2011). Unconscious Thought Theory and its discontents: A critique of the critiques. *Social Cognition*, *29*, 629-647.

Bargh, J. A. & Williams, E. L. (2006). The automaticity of social life. *Current Directions in Psychological Science*, 15, 1-4.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893-912.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230-244.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*, 1014-1027.

Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the Guilty Knowledge Test. *Neuropsychologia*, *22*, 457-469.

Bechara, A., & Damásio, A. R. (2005). The Somatic Marker Hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*, 336-372.

Bechara, A., Damásio, H., Tranel, D., & Damásio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275, 1293-1295.

Bennet, D. & Bennet, A. (2008). Engaging tacit knowledge in support of organizational learning. *VINE: The Journal of Information and Knowledge Management Systems*, 38:1, 72-94.

Binet, A. (1890). On double consciousness: Experimental psychological studies. Chicago: Open Court.

Blair, I. V. & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*, 1142-1163.

Bos, M. W. & Dijksterhuis, A. (2011). Unconscious thought works bottom-up and conscious thought works top-down when forming an impression. *Social Cognition*, *29*, 727-737.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, *81*, 760-773.

Brentano, F. (1874/1973). *Psychology from an empirical standpoint*. A. C. Rancurello,D. B. Terrell, & L. L. McAlister (Trans.). London and New York: Routledge.

Breuer, J. & Freud, S. (1895/1955). *Studies on hysteria*. In *The Complete Psychological Works of Freud. The Standard Edition*. J. Strachey (trans.). Vol. II. London: The Hogarth Press.

Brouwer, A.-M., Zander, T. O., van Erp, J. B. F., Korteling, J. E., & Bronkhorst, A. W. (2015). Using neurophysiological signals that reflect cognitive or affective state: Six recommendations to avoid common pitfalls. *Frontiers in Neuroscience*, *9*. Doi: 10.3389/fnins.2015.00136

Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology*, 40, 227-259.

Budson, A. E. (2009). Understanding memory dysfunction. The Neurologist, 15, 71-79.

Cameron, B. D., Franks, I. M., Enns, J. T., & Chua, R. (2007). Dual-target interference for the 'automatic pilot' in the dorsal stream. *Experimental Brain Research*, *181*, 297-305.

Campion, J., Latto, R., & Smith, Y. M. (1983). Is blindsight an effect of scattered light, spared cortex and near-threshold vision? *Behavioral and Brain Sciences*, *6*, 423-486.

Cardoso-Leite, P. & Gorea, A. (2010). On the perceptual/motor dissociation: A review of concepts, theory, experimental paradigms and data interpretations. *Seeing and Perceiving*, *23*, 89-151.

Casey, P. M., Warren, R. K., Cheesman II, F. L., & Elek, J. K. (2012). *Helping courts address implicit bias. Resources for education*. National Center for State Courts.

Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*, 40-48. Chan, C. (1991). Implicit Cognitive Processes: Theoretical Issues and Applications in Computer Systems Design. Unpublished doctoral dissertation, University of Oxford.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, *36*, 387-395.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, *40*, 343-367.

Cleeremans, A. (1997). Principles for implicit learning. In D. Berry (ed.), *How implicit is implicit learning?* (pp. 195-234). Oxford: Oxford University Press.

Cleeremans, A. & Haynes, D.-J. (1999). Correlating consciousness: A view from empirical science. *Revue Internationale de Philosophie*, *3*, 387-420.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of patternanalyzing skill in amnesia: Dissociation of "knowing how" and "knowing that." *Science*, *210*, 207-210.

Coltheart, M. (2001). Assumptions and methods in cognitive neuropsychology. In B. Rapp (ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. 3-21), New York, NY: Psychology Press.

Coon, D. & Mitterer, J. O. (2010). *Introduction to psychology: Gateways to mind and behavior*. 12th ed. Belmont, CA: Wadsworth.

Cowey, A. (2010). The blindsight saga. Experimental Brain Research, 200, 3-24.

Damásio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Grosset/Putnam.

Damásio, A. R., Tranel, D., & Damásio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, *13*, 89-109.

Damásio, A. R., Tranel, D., Damásio, H. (1991). Somatic markers and the guidance of behavior: Theory and preliminary testing. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (eds.), *Frontal lobe function and dysfunction* (pp. 217–229). New York: Oxford University Press.

Daum, I., Channon, S., & Canavan, A. G. M. (1989). Classical conditioning in patients with severe memory problems. *Journal of Neurology, Neurosurgery, and Psychiatry*, *52*, 47-51.

de Gelder, B., Tamietto, M., van Boxtel, G., Goebel, R., Sahraie, A., van den Stock, J., et al. (2008). Intact navigation skills after bilateral loss of striate cortex. *Current Biology*, *18*, 28-29.

Dehaene, S. (2009). Conscious and nonconscious processes: Distinct forms of evidence accumulation? *Séminaire Poincaré XII*, 89-114.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*, 204-211.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1-37.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752-758.

Demetriou, A., Mouyi, A., Spanoudis, G. (2010). The development of mental processing. In W. F. Overton (ed.), *The handbook of life-span development. Vol. 1: Cognition, biology, and methods* (pp. n.a.), Hoboken, NJ: Wiley.

Descartes, R. (1644/1983). *Principles of philosophy*. V. R. Miller & R. P. Miller, (trans.). Dordrecht, Holland/Boston, USA: Reidel.

Destrebecqz, A. & Peigneux, P. (2005). Methods for studying unconscious learning. *Progress in Brain Research*, *150*, 69-80.

Destrebecqz, A., Franco, A., Bertels, J., & Gaillard, V. (2015). Direct and indirect measures of statistical learning. In M. Overgaard (ed.), *Behavioral methods in consciousness research* (pp. 107-120), Oxford: Oxford University Press.

Devine, P. G. & Sharp, L. B. (2009). Automaticity and control in stereotyping and prejudice. In T. D. Nelson (ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 61-87). New York & Hove: Psychology Press.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. Progress in Brain Research, 168, 49-64.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Sciences*, *6*, 274-290.

Dienes, Z. (2012). Conscious versus unconscious learning of structure. In P. Rebuschat & J. Williams (eds.), *Statistical Learning and Language Acquisition* (pp. 337-364). Mouton de Gruyter.

Dienes, Z., Altmann, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1322-1338.

Dienes, Z. & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, *4*, 3-23.

Dienes, Z., & Perner, J. (2002). A theory of the implicit nature of implicit learning. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, philosophical, and computational consensus in the making?* (pp. 68-92). New York: Psychology Press.

Dienes, Z. & Perner, J. (2004). Assumptions of a subjective measure of consciousness: Three mappings. In R. Gennaro (ed.), *Higher order theories of consciousness* (pp. 173-199). Amsterdam: John Benjamins.

Dijksterhuis, A. (2010). Automaticity and the unconscious. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (eds.), *Handbook of social psychology* (5th ed.), vol. 1 (pp. 228-267). Hoboken, NJ: Wiley.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, *1*, 95-109.

Dijksterhuis, A., Smith, P. K., van Baaren, R. B., & Wigboldus, D. H. J. (2005). The unconscious consumer: Effects of environment on consumer behavior. *Journal of Consumer Psychology*, *15*, 193-202.

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS ONE*, 7. Doi: 0.1371/journal.pone.0029081

Draine, S. C. & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, 127, 286-303.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541-555.

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews*, 30, 239-271.

ECU (2013). Unconscious bias and higher education. London: Equality Challenge Unit.

Egeth, H. E. & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, 48, 269-297.

Ellenberger, H. F. (1970). *The discovery of the unconscious: History and evolution of dynamic psychiatry*. New York: Basic Books.

Ellis, H. D. & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences*, *5*, 149-156.

Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*, 157, 239-248.

Ellis, H. D., Young, A. W., Quayle, A. H., & de Pauw, K. W. (1997). Reduced autonomic responses to faces in Capgras delusion. *Proceedings of the Real Society of London, B, 264*, 1085-1092.

Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychology Review*, 67, 279-300.

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

Evans, J. St. B. T. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223-241.

Fang, F. & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *8*, 1380-1385.

Fazio, R. H. & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297-327.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229-238.

Fendrich, R., Wessinger, C. M., & Gazzaniga, M. S. (1992). Residual vision in a scotoma: Implications for blindsight. *Science*, 258, 1489-1491.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*, 74-147.

Figner, B. & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. In M. Schulte-Mecklenbeck, A. Kühberger, & R. Ranyard (eds.), *A handbook of process tracing methods for decision research* (pp. 163-184). New York, NY: Psychology Press.

Fisk, G. D. & Haase, S. J. (2005). Unconscious perception or not? An evaluation of detection and discrimination as indicators of awareness. *The American Journal of Psychology*, *118*, 183-212.

Fiske, S. T. (1989). Examining the role of intent: Toward understanding its role in stereotyping and prejudice. In J. S. Uleman & J. A. Bargh (eds.), *Unintended thought* (pp. 189-211). New York / London: Guilford Press.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). *Cognitive neuroscience: The biology of the mind*. 2nd ed. London: Norton.

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology*, *50*, 141-150.

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

Gonzáles-Vallejo, C., Lassiter, G. D., Bellezza, F. S., & Lindberg, M. J. (2008). "Save angels perhaps": A critical examination of Unconscious Thought Theory and the deliberation-without-attention effect. *Review of General Psychology*, *12*, 282-296.

Graf, P. & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 501-518.

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 164-178.

Graham, S. & Lowery, B. S. (2004). Priming unconscious racial stereotypes about adolescent offenders. *Law and Human Behavior*, *28*, 483-504.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, selfesteem, and stereotypes. *Psychological Review*, *102*, 4-27.

Greenwald, A. G., Klinger, M. R., & Liu, T. J. (1989). Unconscious processing of dichoptically masked words. *Memory & Cognition*, 17, 35-47.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

Harris, C. H., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PLoS ONE*, 8. Doi: 10.1371/journal.pone.0072467

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.

Hassin, R. R., Uleman, J. S., & Bargh, J. A. (eds.) (2005). *The new unconscious*. New York, NY: Oxford University Press.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews – Neuroscience*, *11*, 523-532

Hodges, J. R., Spatt, J., & Patterson, K. (1999). "What" and "how": Evidence for the dissociation of object knowledge and mechanical problem-solving skills in the human brain. *Proceedings of the National Academy of Sciences U.S.A.*, *96*, 9444-9448.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9*, 1-66.

Huizenga, H. M., Wetzels, R., van Ravenzwaaij, D., & Wagenmakers, E.-J. (2012). Four empirical tests of Unconscious Thought Theory. *Organizational Behavior and Human Decision Processes*, *117*, 332-340.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Jacoby, L. L., Lindsay, D. S., & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist*, *47*, 802-809.

Kahneman, D. & Frederick, S. (2002). Representativeness revisited: attribute substitution in intuitive judgement. In T. Gilovich, D. Griffin, & D. Kahneman (eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge: Cambridge University Press.

Kiefer, M. (2007). Top-down modulation of unconscious 'automatic' processes: A gating framework. *Advances in Cognitive Psychology*, *3*, 289-306.

Kiefer, M. (2012). Executive control over unconscious cognition: Attentional sensitization of unconscious information processing. *Frontiers in Human Neuroscience*, 6. doi: 10.3389/fnhum.2012.00061

Köhler, S., & Moscovitch, M. (1997). Unconscious visual processing in neuropsychological syndromes: A survey of the literature and evaluation of models of consciousness. In M. D. Rugg (ed.), *Cognitive Neuroscience* (pp. 305-373). Hove: Psychology Press.

Krauzlis, R. J., Lovejoy, L. P., & Zénon, A. (2013). Superior colliculus and visual spatial attention. *Annual Review of Neuroscience*, *36*. doi: 10.1146/annurev-neuro-062012-170249.

Lazarus, R. S., & McCleary, R. A. (1951). Autonomic discrimination without awareness: A study of subception. *Psychological Review*, *58*, 113-122.

Locke, J. (1690/1959). *An Essay concerning human understanding* (2 Vols.). New York: Dover Publications.

Logan, G. D. (1989). Automaticity and cognitive control. In J. S. Uleman & J. A. Bargh (eds.), *Unintended thought* (pp. 52-74). New York / London: The Guilford Press.

Martens, U. & Kiefer, M. (2009). Specifying attentional top-down influences on subsequent semantic unconscious processing. *Advances in Cognitive Psychology*, *5*, 56-68.

Martens, U., Ansorge, U., & Kiefer, M. (2011). Controlling the unconscious: Attentional task sets modulates subliminal semantic and visuomotor processes differentially. *Psychological Science*, 22, 282-291.

McAndrews, A., Jones, F. W., McLaren, R. P., & McLaren, I. P. L. (2012). Dissociating expectancy of shock and changes in skin conductance: An investigation of the Perruchet effect using an electrodermal paradigm. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 203-208.

McCleary, R. A., & Lazarus, R. S. (1949). Autonomic discrimination without awareness: An interim report. *Journal of Personality*, *18*, 171-179.

McDougall, W. (1924). Outline of psychology. New York: Charles Scribner's Sons.

Merikle, P. M. (1992). Perception without awareness: Critical issues. *American Psychologist*, 47, 792-795.

Merikle, P. M. & Daneman, M. (1998). Psychological investigations of unconscious perception. *Journal of Consciousness Studies*, *5*, 5-18.

Merikle, P. M. & Reingold, E. M. (1991). Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17, 224-233.*

Miller, J. G. (1939). Discrimination without awareness. *American Journal of Psychology*, *52*, 562-578.

Milner, A. D. (2012). Is visual processing in the dorsal stream accessible to consciousness? *Proceedings of the Royal Society B*. doi: 10.1098/rspb.2011.2663.

Milner, A. D., & Goodale, M. A. (2007). Two visual systems re-viewed. *Neuropsychologia*, 46, 774-785.

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron*, *20*, 445-468.

Mitchell, C. J., De Hower, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183-246.

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

Moskowitz, G. B. (2001). Preconscious control and compensatory cognition. In G. B. Moskowitz (ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 333-358). Mahwah, NJ: Erlbaum.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167-184.

Most, S. B., Sorber, A. V., & Cunningham, J. G. (2007). Auditory Stroop reveals implicit gender associations in adults and children. *Journal of Experimental Social Psychology*, *43*, 287-294.

Naccache, L. (2005). Visual phenomenal consciousness: A neurological guided tour. *Progress in Brain Research*, *150*, 185-195.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Nosek, B. A. & Banaji, M. R. (2001). The Go/No-Go Association Task. Social Cognition, 19, 625-664.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (ed.), *Automatic processes in social thinking and behavior* (pp. 265-292). Hove: Psychology Press.

Ortu, D. (2012). Neuroscientific measures of covert behavior. *Behavioral Analysis*, 35, 75-87.

Overgaard, M. (ed.) (2015). *Behavioral methods in consciousness research*. Oxford: Oxford University Press.

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition. Where is it coming from? Where is it now? Where is it going? In B. Gawronski & B. K. Payne (eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 1-15). New York: Guilford Press.

Peirce, C. S., & Jastrow, J. (1884). On small differences in sensation. *Memoirs of the National Academy of Sciences*, *3*, 73-83.

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gabronwski & B. K. Payne (eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 255-277). New York: Guilford Press.

Perruchet, P. (1985). A pitfall for the expectancy theory of human eyelid conditioning. *Pavlovian Journal of Biological Science*, *20*, 163-170.

Perruchet, P. (2015). Dissociating conscious expectancies from automatic link formation in associative learning: A review on the so-called Perruchet effect. *Journal of Experimental Psychology: Animal Learning and Cognition*, 41, 105-127.

Perruchet, P., Cleeremans, A., & Destrebecqz, A. (2006). Dissociating the effects of automatic activation and explicit expectancy on reaction times in a simple associative learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 955-965.

Posner, M. I. & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 77, 317-327.

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning Memory*, *1*, 217-229.

Reder, L. M., Park, H., & Kieffaber, P. D. (2009). Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychological Bulletin*, *135*, 23-49.

Rees, G. & Frith, C. (2007). Methodologies for identifying the neural correlates of consciousness. In M. Velmans & S. Schneider (eds.), *The Blackwell Companion to Consciousness* (pp. 553-566). Oxford: Blackwell.

Reingold, E. M. (2004). Unconscious perception: Assumptions and interpretive difficulties. *Consciousness and Cognition*, *13*, 117-122.

Reingold, E. M. & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*, 563-575.

Reingold, E. M. & Merikle, P. M. (1990). On the inter-relatedness of theory and measurement in the study of unconscious processes. *Mind & Language*, *5*, 9-28.

Reiss, S. (1980). Pavlovian conditioning and human fear: An expectancy model. *Behavior Therapy*, *11*, 380-396.

Richeson, J. A. & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287-290.

Roediger, H. L. (2003). Reconsidering implicit memory. In J.S. Bowers & C. J. Marsolek (eds.), *Rethinking implicit memory* (pp. 3-18). New York, NY: Oxford University Press.

Sabin, J. A. & Greenwald, A. G. (2012). The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: Pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. *American Journal of Public Health*, *102:5*, 988-995.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501-518.

Schacter, D. L. (1992). Implicit knowledge: New perspectives on unconscious processes. *Proceedings of the National Academy of Sciences U.S.A.*, 89, 11113-11117.

Schank, R. C. (2009). The computer's subconscious. *Applied Artificial Intelligence*, *23*, 186-203.

Schmidt, T. & Vorberg, D. (2006). Criteria for unconscious cognition: Three types of dissociation. *Perception & Psychophysics*, 68, 489-504.

Schmolck, H., Kensinger, E. A., Corkin, S., & Squire, L. R. (2002). Semantic knowledge in patient H.M. and other patients with bilateral medial and lateral temporal lobe lesions. *Hippocampus*, *12*, 520-533.

Schneider, W. & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, *27*, 525-559.

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

Seger, C. A., Prabhakaran, V., Poldrack, R. A., & Gabrieli, J. D. E. (2000). Neural activity differs between explicit and implicit learning of artificial grammar strings: An fMRI study. *Psychobiology*, 283, 283-292.

Seth, A.K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2010). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*, 314-321.

Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. Goldstone (eds.), *Handbook of cognition* (pp. 202-220). London: Sage.

Shanks, D. R. & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Brain and Behavioral Sciences*, *17*, 367-447.

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

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-190.

Shimamura, A. P. & Squire, L. R. (1989). Impaired priming of new associations in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 721-728.

Sidis, B. (1898). *The psychology of suggestion*. A research into the subconscious nature of man and society. New York: D. Appleton.

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M-M., & Reber, P. J. (2002). Neural correlates of artificial grammar learning. *NeuroImage*, *17*, 1306-1314.

Spruyt, A. & Tibboel, H. (2015). On the automaticity of the evaluative priming effect in the valent/non-valent categorization task. *PLoS ONE*, *10*. doi: 10.1371/journal.pone.0121564

Squire, L. R. & Wixted, J. T. (2011). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, *34*, 258-288.

Srinivasan, N. & Mukherjee, S. (2010). Attribute preference and selection in multiattribute decision making: Implications for unconscious and conscious thought. *Consciousness & Cognition*, *19*, 644-652.

Stanovich, K. E. (1999). Who is rational? Studies of individual differences in reasoning.Mahwah, NJ: Erlbaum.

Steele, J. R. & Ambadi, N. (2006). "Math is hard!"The effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology*, 42, 428-436.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.

Sun, R. (1999). Accounting for the computational basis of consciousness: A connectionist approach. *Consciousness & Cognition*, *8*, 529-565.

Sun, R. (2002). *Duality of the mind: A bottom up approach toward cognition*. Mahwah, NJ: Erlbaum Associates.

Sun, R. (2008). Introduction to computational cognitive modeling. In R. Sun (ed.), *The Cambridge handbook of computational psychology* (pp. 1-19), Cambridge, etc.: Cambridge University Press.

Sun, R., Lane, S. M., & Mathews, R. C. (2009). The two systems of learning: An architectural perspective. In J. Evans & K. Frankish (eds.), *Two minds: Dual processes and beyond* (pp. 239-264). New York: Oxford University Press.

Sun, R., Mathews, R. & Lane, S. M. (2007). Implicit and explicit processes in the development of cognitive skills: A theoretical interpretation with some practical implications for science education. In E. Vargios (ed.), *Educational Psychology Research Focus* (pp. 1-26), Hauppauge, NY: Nova Science Publishers.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: a bottom-up model of skill learning. *Cognitive Science*, *25*, 203-244.

Sun, R. & Zhang, X. (2004). Top-down versus bottom-up learning in cognitive skill acquisition. *Cognitive Systems Research*, *5*, 63-89.

Swift, J. D. (2003). Mechanisms of memory. San Diego, CA: Elsevier.

Taffler, R. J. & Tuckett, D. A. (2010). Emotional finance: The role of the unconscious in financial decisions. In H. K. Baker & J. R. Nofsinger (eds.), *Behavioral finance. Investors, corporations, and markets* (pp. 95-112), Hoboken, NJ: John Wiley & Sons.

Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*, 697-709.

Thagard, P. (2005). *Mind: Introduction to cognitive science*. 2nd ed. Cambridge, MA, & London: MIT Press.

Thagard, P. (2012). Cognitive architectures. In K. Frankish & W. M. Ramsey (eds.), *The Cambridge handbook of cognitive science* (pp. 50-70), Cambridge, etc: Cambridge University Press.

Timmermans, B. & Cleeremans, A. (2015). How can we measure awareness? An overview of current methods. In M. Overgaard (ed.), *Behavioral methods in consciousness research* (pp. 21-46), Oxford: Oxford University Press.

Tranel, D., Damásio, H., & Damásio, A. R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience*, *7*, 425-432.

Tranel, D., Damásio, A. R., Damásio, H., & Brandt, J. P. (1994). Sensorimotor skill learning in amnesia: Additional evidence for the neural basis of nondeclarative memory. *Learning & Memory*, *1*, 165-179.

Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, *14*, 411-443.

Tulving, E., Hayman, C. A. G., & Macdonald, C. A. (1991). Long-lasting perceptual priming and semantic learning in amnesia: A case experiment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 595-617.

Uleman, J. S. & Bargh, J. A. (eds.) (1989). *Unintended thought*. New York / London: Guilford Press.

Umbach, V. J., Schwager, S., Frensch, P. A., & Gaschler, R. (2012). Does explicit expectation really affect preparation? *Frontiers in Psychology*, *3*, 1-12.

Van Opstal, F., Calderon, C. B., Gevers, W., & Verguts, T. (2011). Setting the stage subliminally: Unconscious context effects. *Consciousness & Cognition*, 20, 1860-1864.

VanRullen, R. & Koch, C. (2005). Visual attention and visual awareness. In G. G. Celesia (ed.), *Disorders of visual processing: Handbook of clinical neurophysiology Vol. 5* (pp. 65-83). Elsevier.

Weidemann, G., Tangen, J. M., Lovibond, P. F., & Mitchell, C. J. (2009). Is Perruchet's dissociation between eyeblink conditioned responding and outcome expectancy evidence for two learning systems? *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 169-176.

Weiskrantz, L. (1986). *Blindsight: A case study and implications*. Oxford: Clarendon Press.

Weiskrantz, L. & Warrington, E. K. (1979). Conditioning in amnesic patients. *Neuropsychologia*, *17*, 187-194.

Willingham, D. B. & Preuss, L. (1995). The death of implicit memory. Psyche, 2, 1-10.

Ziori, E. & Dienes, Z. (2008). How does prior knowledge affect implicit and explicit concept learning? *Quarterly Journal of Experimental Psychology*, *61*, 601-624.

66