In two minds: dual-process accounts of reasoning

Jonathan St. B.T. Evans

Centre for Thinking and Language, University of Plymouth, Plymouth, PL4 8AA, UK

Researchers in thinking and reasoning have proposed recently that there are two distinct cognitive systems underlying reasoning. System 1 is old in evolutionary terms and shared with other animals: it comprises a set of autonomous subsystems that include both innate input modules and domain-specific knowledge acquired by a domain-general learning mechanism. System 2 is evolutionarily recent and distinctively human: it permits abstract reasoning and hypothetical thinking, but is constrained by working memory capacity and correlated with measures of general intelligence. These theories essentially posit two minds in one brain with a range of experimental psychological evidence showing that the two systems compete for control of our inferences and actions.

The idea that there are two distinct kinds of reasoning has been around for as long as philosophers and psychologists have written about the nature of human thought. However, it is only in recent years that cognitive scientists have proposed the striking and strong claim that there are two quite separate cognitive systems underlying thinking and reasoning with distinct evolutionary histories. These two systems are sometimes described as Implicit and Explicit [1,2] although some dual-process theorists prefer to emphasize the functional differences between the two systems and leave open the relation to consciousness [3,4]. In this article, I will therefore use the neutral terms System 1 and System 2 as introduced by Stanovich and West [5,6]. Contemporary interest in the dual-process accounts of reasoning is evidenced by the wider application to related fields such as judgment and decision making [7] and exciting developments in neuropsychological studies of reasoning, described below.

System 1 is generally described as a form of universal cognition shared between humans and animals. It is actually not really a single system, but a set of sub-systems that operate with some autonomy [8,9]. System 1 includes instinctive behaviours that are innately programmed, and would include any innate input modules of the kind proposed by Fodor [10] which are not be confused with more questionable [11] recent claims for domain-encapsulated innate modules that control specific behaviours. The System 1 processes that are most often described, however, are those that are formed by associative learning processes of the kind produced by neural networks [12]. The autonomy of such systems reflects the domain-specific nature of the learning, even though the learning mechanism itself is domain-general [13]. Dual-process theorists generally agree that System 1 processes are rapid, parallel and automatic in nature: only their final product is posted in consciousness. There is at least one contemporary research programme in which researchers are attempting to account for all reasoning results in terms of System 1 level processes [14]. However, I shall provide substantial evidence that postulation of a second system is required.

System 2 is believed to have evolved much more recently and is thought by most theorists to be uniquely human. System 2 thinking is slow and sequential in nature and makes use of the central working memory system that has been so intensively studied in the psychology of memory [15,16]. Despite its limited capacity and slower speed of operation, System 2 permits abstract hypothetical thinking that cannot be achieved by System 1. Consider the case of decision-making. We might (and frequently do) decide our actions on the basis of past experience, doing what has worked well in the past. Such intuitive decisions require little reflection. However, we can also make decisions by constructing mental models or simulations of future possibilities, a process that I term ‘hypothetical thinking’. This distinctively human facility – provided by System 2 – is of great importance. We cannot, for example, learn by experience to avoid disasters such as nuclear war or the effects of uncontrolled global warming.

I will focus in this review principally on the theories that have arisen in the cognitive study of reasoning. However, these theories have been complemented by and in some cases influenced by similar developments in related fields, notably in the study of implicit and explicit learning [2,17], conceptual thinking and categorization [4], social judgment theory [18] and cognitive social psychology [19]. We should also note (see below) that reasoning researchers are increasingly emphasizing the inhibitory role of System 2 in suppressing default knowledge and belief based responses, which provides clear links with contemporary research on the inhibitory role of executive processes in the study of working memory [16].

Evidence for dual process in reasoning

In this review, I discuss the development of dual-process accounts of reasoning and the contribution that reasoning researchers have made to the development of theoretical assessment of the underlying mechanisms. The majority of these studies have used the deductive reasoning paradigm...
Box 1. The belief-bias effect

One of the key methods for demonstrating dual processes in reasoning tasks involves the so-called ‘belief-bias’ effect. The methodology introduced by Evans et al. [66] seeks to create a conflict between responses based upon a process of logical reasoning and those derived from prior belief about the truth of conclusions. Typically, syllogisms are presented for evaluation, which fall into one of the four following categories:

1. Valid argument, believable conclusion (NO CONFLICT)
   Example:
   
   No police dogs are vicious
   Some highly trained dogs are vicious
   Therefore, some highly trained dogs are not police dogs

2. Valid argument, unbelievable conclusion (CONFLICT)
   Example:
   
   No nutritional things are inexpensive
   Some vitamin tablets are inexpensive
   Therefore, some vitamin tablets are not nutritional

3. Invalid argument, believable conclusion (CONFLICT)
   Example:
   
   No addictive things are inexpensive
   Some cigarettes are inexpensive
   Therefore, some addictive things are not cigarettes

4. Invalid argument, unbelievable conclusion (NO CONFLICT)
   Example:
   
   No millionaires are hard workers
   Some rich people are hard workers
   Therefore, some millionaires are not rich people

In belief-bias experiments, participants are instructed to treat the problem as a logical reasoning task and to endorse only conclusions that necessarily follow from the premises given. In spite of this, intelligent adult populations (undergraduate students) are consistently influenced by the prior believability of the conclusion given as well as by the validity of the arguments presented. The conclusion endorsement rates from the study of Evans et al. [66] are shown in Figure I. It is clear that participants are substantially influenced by both the logic of the argument and believability of its conclusion, with more belief-bias on invalid arguments. Dual-process accounts propose that although participants attempt to reason logically in accord with the instructions, the influence of prior beliefs is extremely difficult to suppress and effectively competes for control of the responses made.

Box 2. Neuropsychological evidence for dual processes in reasoning

Vinod Goel and his colleagues have recently demonstrated evidence for dual processes using neuropsychological methods. In one study, using fMRI methodology, they showed evidence for neural differentiation of reasoning with abstract materials and with semantically rich problem materials [27]. Content based reasoning recruited a left hemisphere temporal system whereas reasoning with abstract formal problems was associated with activation of a parietal system. The authors concluded that reasoning was implemented in two systems, depending upon the use of semantic content.

More specifically informative and interesting is the recent study of Goel and Dolan [28]. In this study they used the belief-bias paradigm described in Box 1. Specifically, they were interested in the conflict problems: valid–unbelievable and invalid–believable. Study of belief-logic conflict problems provides one of the key pieces of evidence for dual-process theories of reasoning. In particular it supports the idea that System 2 process can intervene or inhibit System 1 processes which will otherwise lead to pragmatic or belief-based responding on a task where deductive reasoning is required by the instructions presented.

Goel and Dolan, again using fMRI techniques, discovered that the resolution of such conflict problems in favour of either logic or belief was differentiated with respect to associated neurological activity. Specifically, on trials where the logically correct decision was made, responses were associated with activation of the right inferior prefrontal cortex. By contrast, incorrect, belief-biased responses were associated with activation of the ventral medial prefrontal cortex. This provides very strong evidence for long standing claims of dual-process theorists that different mental processes are competing for control of the response to these problems [66]. Goel and Dolan conjecture that the right prefrontal cortex is critical in detecting and resolving conflict, a key aspect of System 2 functioning, also supported in their earlier study [27]. They also cite a range of neuropsychological studies to support their contention that the ventral medial prefrontal cortex is associated with a range of intuitive or heuristic responses of the kind typically characterised as emanating from System 1.

In which participants untrained in formal logic are asked to judge the validity of arguments, basing their responses only on the information provided [20]. The paradigm has its origins in logicism – the view that logic provides the basis for rational thinking. Logicist thinking has, however, been progressively undermined in this research field because of numerous demonstrations that problem content and context affect the way in which people reason on such tasks. Most people find it very hard to disregard the meaning of the material with which they are asked to reason.

One of the key findings in this literature is that of a ‘belief-bias’ effect (Box 1). Student participants are instructed to assess the logical validity of arguments whose conclusions are a priori believable or unbelievable. It appears that both logical and belief-based processes are influencing the task and may be in competition with one another. In the dual-process account, these are attributed to Systems 2 and 1 respectively. Of particular importance are problems that bring belief and logic into conflict (see Box 1). The ability to resolve such conflict in favour of logic is known to be correlated with measures of general cognitive ability [21] and to decline sharply with age [22]. Recent experimental studies have enhanced our understanding of these effects [23,24].

Very strong instructional emphasis on logical necessity
Box 3. The Wason selection task

The Wason selection task provides important evidence for dual-process accounts of reasoning, because performance with the task is so sensitive to the content and context with which it is presented. The abstract, indicative selection task is illustrated in Figure Ia. The correct answer is generally agreed to be A and 7, although this has been disputed by some theorists [67,68]. The statement can only be falsified by finding a case of a card that has an A on one side and does not have a 3 on the other. Only turning the A and 7 (not a 3) can lead to discovery of such a case. However, few people (10–20%) give this correct answer when tested.

Performance on the abstract selection task is thought to be strongly influenced by a System 1 heuristic known as ‘matching bias’ [30]. The effect is demonstrated by introducing negative components into the conditional statement. Imagine that the rule presented in Figure Ia was ‘If there is an A one side of the card, then there is not a 3 on the other side’. The A and 3 cards, which have a perceptual match to the items named in the rule, are now also the logically correct cards, because discovery of an A3 card would falsify the rule. There is hence a strong tendency to choose matching cards, regardless of their logical status. This is one of the most reliable and robust biases in the psychological reasoning.

When the difficult affirmative form of the Wason selection task is couched in realistic terms it is very much easier to solve. For example, Griggs and Cox [69] reported around 75% successful solutions of a problem in which they were instructed to imagine that they were police officers observing people drinking in a bar (Figure Ib). Most participants selected the cards showing ‘Drinking beer’ and ‘16 years of age’, the equivalent of the A and 7 cards on the abstract task described above. These choices could lead to discovery of an under-age drinker who is breaking the rule. This kind of problem is often described as a deontic selection task [70], as it concerns following rules rather than truth and falsity.

**Figure I.** (a) in the abstract, indicative version of the Wason selection task, participants are given a conditional statement and shown four cards, each of which is known to have a letter on one side and a number on the other. The task is to decide which cards need to be turned over to find out whether the statement is true or false. In the following example, typical choices are A or A and 3, but the logically correct choices is A and 7. (b) In a realistic and deontic version of the selection task, the participants are given a rule or regulation and need to check whether it is being obeyed. In the following example, they play a police officer checking people drinking in a bar. The cards represent different drinkers with the beverage on one side and the age on the other. Most people correctly choose the beer drinker and the 16 year old.

Individual differences in reasoning ability

System 2 requires working memory whose capacity is known to vary across individuals. In fact, working memory capacity and reasoning ability are known to be highly correlated [33–35]. Consistent with this, it has been argued by dual-process theorists [2,26] that System 2 function should be related to measures of general intelligence, although System 1 function should be independent of such measures. Stanovich and West have
demonstrated in a recent series of studies [6,26,36,37] that the ability of participants to find normatively correct solutions to a range of inferential and decision making tasks was consistently associated with those who were high in cognitive ability as measured by SAT scores. However, they have also found important exceptions which accord with predictions of the dual-process theory.

One of the most investigated reasoning problems in the literature is the Wason selection task (see Box 3). In an abstract, indicative version (Figure 1a in Box 3), it is known to be very difficult. However, in a realistic, deontic version (Figure 1b), is it quite easy. On the dual-process account, the former is difficult because it requires explicit and abstract logical reasoning of the kind that only System 2 can provide. With the latter task, however, the correct answer is strongly cued by relevant prior knowledge, reflecting System 1 processes [1]. Stanovich and West [38] reasoned that on this basis, solution of the abstract, but not the thematic task should be sensitive to measure of cognitive ability. Accordingly, they reported much sharper differences in SAT scores between solvers and non-solvers of the abstract problem. (For recent debate about the nature of domain-specific cognitive processes on the deontic selection task, see [39,40].)

It is important to note that it is abstract reasoning and the ability to comply with instructions, rather than logical reasoning as such, that differentiates those high in cognitive ability and System 2 functioning. As Stanovich and West’s work also shows, high ability participants can also solve statistical and decision making problems, and are better able to resist the contextualisation of problems within prior knowledge and belief. Further support for their conclusions comes from recent applications of dual-process theory to the development of reasoning in children and adolescents [41,42]. System 2 processing is more strongly linked than System 1 processing, to the child’s age and measured intelligence. These findings are further complemented by evidence that System 2 function declines, relative to System 1, in old age [22].

Although System 2 processing is clearly implicated in any account of the logical competence that ordinary people can exhibit in deductive reasoning, it should be noted that the dual-process theory does not take sides on the issue of whether this competence is achieved by manipulation of mental models [43,44] or mental rules [45,46]. Indeed, it can be argued that both theories involve implicit distinctions that map on to the dual-process account [47]. Theorists of both persuasions have proposed mechanisms that account of deductive competence (System 2) and additional mechanisms that provide an account of pragmatic influences.

### Dual processes in judgment and decision-making

Although the dual-process theories discussed in this review were largely stimulated by research on reasoning, they should be equally relevant to studies of judgment and decision making, as has recently been recognized by researchers in these fields [7,48–51]. Research in the ‘heuristics and biases’ tradition, stimulated by a series of seminal papers by Amos Tversky and Daniel Kahneman [52], traditionally emphasized the role of short-cut heuristics in probability judgment and the cognitive biases that resulted. Application of a similar approach to the study of reasoning [53] was a precursor of the dual-process theory. Just as reasoning theorists came to understand that unconscious biases could be overridden by an explicit effort at reasoning, so judgment researchers have recently reached a similar conclusion. Kahneman and Frederick stated, ‘We assume that System 1 and System 2 can be active concurrently, that automatic and controlled cognitive operations compete for the control of overt responses...’ [51], pp. 51–52).

Recent studies of Bayesian reasoning have been focussed on arguments about whether the phenomena are best accounted for in terms of an innate frequency processing module associated with natural sampling [54–56] – a System 1 level of explanation, or whether the experimental procedures that facilitate statistical reasoning actually support explicit reasoning in System 2 [57–59]. There is also an apparent conflict between the finding that people are biased by prior beliefs in deductive reasoning, but neglectful of prior probabilities in Bayesian reasoning. Recently, Evans et al. [60] argued that this was because the great majority of studies presented base rates as explicit statistical information, requiring difficult System 2 reasoning to compute their effect on posterior probabilities. They showed that people were much more likely to take account of base rates provided by real-world prior beliefs of the participants, thus presumably recruiting the assistance of System 1 processes. This finding accords with those of studies showing that base rates acquired by experimental training are more likely to be used [61].

### Dual-process theory and evolutionary arguments

Dual-process theorists claim that human beings evolved a powerful general purpose reasoning system – System 2 – at quite a late stage, and this co-exists with a much older set of autonomous sub-systems labelled as System 1. There is evidence in the archaeological record that lends credibility to this claim [62]. In commenting on the remarkable competitive success of Homo Sapiens Sapiens over other hominids, Mithen stated, ‘This persuades many archaeologists that modern humans had…a cognitive advantage which may have resided in a more complex form of language or a quite different type of mentality... Support for the latter is readily evident in from dramatic developments that occur in the archaeological record relating to new ways of thinking and behaving by modern humans.’ [62], p. 33). Mithen comments on the qualitative change in the archaeological record c.50,000 years ago when there was sudden evidence of representational art, religious imagery and rapid adaptations in the design of tools and artefacts.

Dual-process theorists find themselves in some conflict with evolutionary psychologists who claim that the mind is massively modular [63], or who downplay the role of general reasoning ability in favour of domain-specific mechanisms [39]. Such approaches appear greatly to underestimate the role of System 2 processes [8,13,64]. Of course, it is unparsimonious to propose a specific mechanism for a process that can be accounted for by a
general one. It has also been argued that apparently domain-specific reasoning can be accounted for by pragmatic principles from relevance theory, that are in fact quite general in operation [65]. This debate has linked closely with another about the nature of human rationality. Stanovich and West [6,8] have argued for a distinction between evolutionary and individual rationality. The idea is that once System 2 evolved, it became a 'long-leash' system with little direct genetic control, allowing humans to pursue their own individual goals rather than to act merely as slavish vehicles of the genes [9].

Summary and conclusions

Dual-process theories of thinking and reasoning quite literally propose the presence of two minds in one brain. The stream of consciousness that broadly corresponds to System 2 thinking is massively supplemented by a whole set of autonomous subsystems in System 1 that post only their final products into consciousness and compete directly for control of our inferences, decisions and actions. However, System 2 provides the basis for hypothetical thinking that endows modern humans with unique potential for a higher level of rationality in their reasoning and decision-making.

There are several important directions for future research. Current theories are framed in general terms and are yet to be developed in terms of their specific computational architecture. An important challenge is to develop models to show how such two distinct systems interact in one brain and to consider specifically how the conflict and competition between the two systems might be resolved in the control of behaviour. Although early results are encouraging, neuropsychological studies of reasoning are in their infancy and substantial research effort will be needed to develop our understanding of the neurological basis of the dual systems. Theoretical and experimental psychologists need to focus on the interaction of the two systems and the extent to which volitional process in System 2 can be used to inhibit the strong pragmatic tendencies to respond in inference and judgment that come from System1, especially where the latter are known to result in cognitive biases. An effort is needed to relate this work on reasoning to the highly relevant, but then and now. In Heuristics and Biases: The Psychology of Intuitive Judgement (Gilovich, T. et al., eds.), pp. 1–18, Cambridge University Press

References