Perspectives on Psychological Science
8(3) 248–252

DOI: 10.1177/1745691613483475
pps.sagepub.com

A Case Study: Dual-Process Theories of Higher Cognition—Commentary on Evans & Stanovich (2013)

Magda Osman
Queen Mary, University of London, England

Abstract
Dual-process theories of higher order cognition (DPTs) have been enjoying much success, particularly since Kahneman's 2002 Nobel prize address and recent book *Thinking, Fast and Slow* (2009). Historically, DPTs have attempted to provide a conceptual framework that helps classify and predict differences in patterns of behavior found under some circumstances and not others in a host of reasoning, judgment, and decision-making tasks. As evidence has changed and techniques for examining behavior have moved on, so too have DPTs. Killing two birds with one stone, Evans and Stanovich (2013, this issue) respond to five main criticisms of DPTs. Along with addressing each criticism in turn, they set out to clarify the essential defining characteristics that distinguish one form of higher order cognition from the other. The aim of this commentary is to consider the defining characteristics of Type 1 and Type 2 processing that have been proposed and to suggest that the evidence can be taken to support quantitative differences rather than qualitatively distinct processes.

Introduction

If one were an inexperienced researcher joining the study of dual-process theories of higher cognition (DPTs) for the first time, then there might be a few questions that might come to mind. What is the fundamental difference (or differences) that distinguishes one process of higher cognition (e.g., Type 1 reasoning) from the other (e.g., Type 2 reasoning)? Is this fundamental difference (or differences) preserved in other forms of higher cognition such as judgment, decision making, and perhaps even problem solving? How can DPTs be falsified?

Evans and Stanovich's (2013, this issue; hereafter, E&S) careful responses and clarification of DPTs in effect give direct answers to these research questions. To answer the first question, through E&S's exposition, we as researchers into higher order cognition (e.g., inductive, deductive causal reasoning, judgment, and decision making) can now refer to both necessary and sufficient attributes of Type 1 processing [T1] and Type 2 processing [T2]. One of the necessary defining features of T1 and T2 processing is the differential loading on working memory (WM): There is no loading for T1 and some loading for T2. The other necessary defining feature of T1 is that it is autonomous (which is not the same as being fast), and another defining feature of T2 is that it enables mental simulation. Moreover, E&S have now brought some clarity to an increasingly detailed and ever changing list of features that have been associated with T1 and T2 processing. In response to the second question, E&S are clear in that they propose that T1 and T2 are found in different forms of higher cognition which include reasoning, judgment and decision making. Finally, in response to the third question, given the necessary and sufficient criteria for identifying different processes of higher cognition, it is possible to imagine ways of setting up an experimental procedure that could attempt to falsify a hypothesis originating from a DPT position. More to the point, it is clearer now what evidence would appear to challenge DPTs.

Given E&S's illuminating article, the aim of this commentary is to explore Criticism 5: "Evidence for dual processing is ambiguous or unconvincing" (p. 232). In order to do this, the discussion that follows considers two defining characteristics of T1 and T2 processing with

Corresponding Author:
Magda Osman, Centre for Biological and Experimental Psychology, School of Biological and Chemical Sciences, Queen Mary, University of London, Mile End Road, London, E1 4NS, England
E-mail: m.osman@qmul.ac.uk

respect to findings that E&S themselves cite in support of their DPT position: working memory capacity (WMC) and automaticity (De Neys, 2006a, 2006b). The commentary begins by showing that evidence judged to support qualitatively distinct processes can be reinterpreted as evidence for quantitatively different processes and that definitions of automaticity require further refinement to avoid continued ambiguity (Osman, 2004). The second section raises a question about hypothetical thinking in causal reasoning and whether it should be classified as T2 processing. Finally, the conclusion considers the future of DPTs.

Supporting Evidence for DPTs?

Let's take De Neys's (2006b) article as our first case study for considering the necessary defining distinctions between T1 and T2 processing. In the study, all participants were first classified into three groups (low, medium and high) based on their WMC. WMC was assessed by their performance on a WM task; this was a dual-processing task based on accurate recall of unrelated words while performing mathematical operations. The main part of the experiment involved a manipulation of the syllogisms that participants were evaluating. The believability of the conclusion of each syllogism was either incongruent with the logical conclusion (conflict syllogism) or congruent with the logical conclusion (no-conflict syllogism). A second manipulation concerned WM loading (high, low, none) while evaluating the syllogisms. The load was varied by using a *dot memory task*, which involved the presentation of an array of dots in a 3 × 3 grid with each reasoning task sandwiched between a dot task; people saw the array before each reasoning task, and their job was to replicate the array after evaluating the syllogism.

Reinterpreting evidence of qualitative differences between T1 and T2 processes as quantitative differences based on the difficulty of the task

Let's start with a simple prediction that T1 processing is found in the no-conflict syllogism, because the syllogisms are easy to evaluate correctly and thus there should be no effect of WM load. Why might they be easy to evaluate correctly? When logic and belief agree, there is a tendency to accept the syllogism as valid (belief bias effect; Wilkins, 1928), and the premises and conclusion are easily integrated and evaluated because they cohere. This can facilitate what seems like a logical response when actually people are simply going by how believable the conclusion is: believable = valid, and unbelievable = invalid. De Neys (2006b) did indeed find support for this prediction. Regardless of WMC and WM load, accuracy was near to ceiling. So the next most obvious question to ask is "Did people use T1 or T2 processing to solve the no-conflict tasks?". Given that there was no loading on WM, then presumably responses were generated by T1. Why? Because relying on real world knowledge to evaluate a syllogism is likely to be easier (and will not require extensive cognitive resources) than implementing the rules of logic to assess the syllogism's validity. That is to say that, given the near ceiling performance and lack of WM loading, all people responding to no-conflict tasks in De Neys's (2006b) study were more likely to be using a T1 process when responding.

What about in the conflict syllogism task? Again, let's adopt a simple prediction that T2 processing is found in conflict tasks, because the syllogisms are difficult to evaluate and, thus, there should be an effect of WM load. Why might they be difficult to evaluate? Because when logic and belief disagree, people will tend to base their responses on the believability of the conclusion, but they have a harder time integrating the premises and conclusion because the result of the integration challenges their prior knowledge and experience of relations in the world (Klauer, Musch, & Naumer, 2000). De Neys (2006b) also found support for this prediction. In fact, he also reported similar findings in an earlier study (De Neys & Dieussaert, 2005). Regardless of the WMC of the participants evaluating the syllogisms, accuracy of performance was severely affected as the load increased. Given that loading on WM affected performance regardless of WMC, it appears that responses were generated by T2. That is to say that all people responding to conflict tasks in De Neys's (2006b) study were responding using a T2 process, because all of them showed an effect of loading on WM while solving the conflict syllogisms.

The difference in results between conflict and no-conflict syllogisms fits perfectly well with a basic prediction that would derive from E&S's DPTs. Easy tasks do not load on WM and do not affect accuracy, thus they are performed using T1; whereas hard tasks load on WM and affect accuracy, thus they are performed using T2. However, the interpretation is based on a quantitative distinction on a single dimension, which is the difficulty of the task. Of course, the rather simplistic prediction that I have proposed for interpreting the actual data that De Neys finds is not quite what E&S's DPTs and other DPTs, (e.g., Epstein, 1994; Evans, 2003; Evans & Over, 1996; Sloman, 1996; Stanovich & West, 2000) would have in mind.

What doesn't sit well with them is that WMC, which would typically be used by dual processing to demonstrate individual difference in performance, should actually be shown to reveal fundamental qualitative differences between T1 and T2—not merely quantitative differences. In the case of conflict syllogisms, a DPT sympathizer would predict the following: Those with low WMC should perform worse than those with high WMC—De Neys (2006b) did find this. Next, a DPT sympathizer would predict that those with low WMC should not show an effect of loading on WM because they are using automatic processes to solve the task, where as those with high WMC should show an effect of loading on WM because they are using analytic processes that are typically slow and drain WMC. In effect, DPTs would predict an interaction between WMC and WM load, which De Neys (2006b) did not find. Therefore, De Neys's (2006b) study doesn't support a qualitative distinction between T1 and T2. Instead, the findings can be taken to show that the same reasoning process generates accurate responses to easier tasks and generates inaccurate responses to harder tasks.

Reinterpreting evidence of qualitative differences between T1 and T2 based on automaticity as quantitative differences

E&S's DPTs are explicit about the fact that only differential loadings on WM and autonomous and hypothetical features definitively distinguish T1 processing from T2 processing, and speed of processing is a sufficient but not necessary feature. Given the description of the relationship between T1 and T2 processing, T1 should typically generate responses faster than T2, especially since the collection of phenomena under T1 processing are autonomous, and processes that are automatic tend to be fast (Posner, 1978). We can explore the autonomous nature of T1 processing and its immunity from loading on WM in De Neys's (2006a) findings, which E&S take as support for their reformulation of DPTs.

In De Neys's (2006a) study, a variety of well-worn reasoning and judgment tasks were used (e.g., the "Linda problem" and the deontic and indicative versions of the selection task) in combination with a variety of methods of loading on WM. For reasons of complementarity with the previous case study, I will focus on the findings from the selection task because the WM loading manipulation was the same as De Neys (2006b) study. Though I should note that I have not carefully selected findings to suit my purposes, a reading of De Neys's (2006a) study will show that I have accurately reported the headline results (also, see De Neys, 2012). The selection task involves a set of four cards that are presented along with a conditional statement such as "if there is a vowel on one side, then there is an even number on the other" (indicative—a context that refers to an arbitrary conditional rule) or "If a person is drinking beer, then the person needs to be over 21 years of age" (deontic—a context that refers to regulations). People are asked to select cards that would test the rule. One uncontroversial point that can be made is that the main reason for the boost in performance in deontic tasks (as compared with indicative tasks) is that prior knowledge about real world infractions of regulations facilitates responses. Moreover, deontic tasks are hypothesized to be performed using T1 processing, which is automatic and fast by definition (Evans, 2003; Evans & Over, 1996; Sloman, 1996; Stanovich & West, 2000).

De Neys (2006a) examined the effects of WM loading on accuracy of performance in deontic and indicative tasks. A simple DPT prediction would be that there should be no effect of WM load on accuracy in deontic tasks, but there should be an effect of WM load on accuracy in indicative versions. De Neys (2006a, Experiment 2) showed that, in both Deontic and Indicative versions of the selection task, accuracy of performance was poorer when there was a load on WM than when there was no load—a finding also reported by Oaksford, Morris, Grainger, and Williams (1996). Thus, the prediction was not supported. Instead, the findings tell us that reasoning in deontic and indicative tasks requires WM. The more WM is loaded, the worse performance is in both versions. This means that either deontic tasks are not solved by T1 processes (which are automatic) or that T1 processes are not automatic. More to the point, the findings clearly contradict E&S's DPTs and most other DP accounts.

How automatic is automatic in the case of T1 processing?

Surprisingly, there are few studies that have actually recorded participants' response latencies while reasoning (De Neys, 2006a; Osman, 2007; Roberts & Newton, 2001) in order to examine speed of processing of T1 and T2. Other than its immunity from loading on WM, the discussion now focuses on T1 processing as a fast process. As mentioned previously, T1 processing is commonly associated with deontic tasks, so one question that can be posed is "How fast is reasoning on Deontic tasks?". To answer this, we can return to De Neys (2006a). As well as manipulating load for deontic and indicative tasks, De Neys recorded latencies for time spent reading the instructions and time spent making inferences. For deontic tasks, people took about 26 s to make an inference under load conditions (versus 20 s for indicative tasks) and they spent about 22 s making an inference without a load (versus 20 s for indicative tasks). When time spent making inferences was compared for both versions, regardless of load manipulation, there was no difference (De Neys, 2006a; Roberts & Newton, 2001). This result is problematic because deontic tasks are supposedly solved by T1 processing, and T1 is automatic, so responses should at least be faster for deontic tasks than for indicative tasks. Also, if T1 is immune from WM loading, then responses should be faster for deontic tasks. Neither of these claims was supported by the evidence.

Is it automatic if one takes 26 s to make an inference based on prior knowledge? Another problem that the evidence raises is that there is no clear definition of what automatic T1 processing is. Stanovich (2004) proposed TASS (the autonomous set of systems), which refers to a set of processes that fall under what E&S's DPTs call T1 and respond automatically to triggering stimuli. If we stay in the domains of higher order cognition, then we can try to examine the speed of processes that are part of TASS by considering the Linda problem. This task is an example in which erroneous responses are thought to be generated by T1. The reason for this is that the information presented in the task automatically triggers prior knowledge used to help make a judgment about the likelihood of various statements corresponding to a fictional character called Linda that conflict with the correct response. Fortunately De Neys (2006a) recorded judgment times. People spent about 47 s making the erroneous judgments and about 57 s making the correct judgments. Clearly making the right responses takes longer, but the important question here is whether 47 s reflects an autonomous response: 47 s appears to be rather a long time (excluding time taken to read the instructions) and certainly more than the 19 s (on average) taken to make the correct response in the deontic version of the selection task (De Neys, 2006a, Experiment 1). This is a crude comparison, but hopefully it makes the point. Of the collection of tasks that are automatic, the response times are not uniformly the same—they need not be. However, there does need to be an agreed point of comparison. The term *automatic* can only have resonance if it precisely specified relative to a specific benchmark. What this would require is a clear definition of the contribution of T1 processing and T2 processing in time spent reading instructions, interpreting instructions, encoding task information, evaluating the information, and making the response. More to the point it would also require specifying where exactly in these various activities automaticity comes in to play. This, at least, is what goes on in other domains of cognition in which automaticity in processing is studied (e.g., Hommel, 2000).

Warnings From Other Domains, Some Reflections and Questions

My final direct comment on the characteristics defining T1 and T2 concerns the association between WMC and hypothetical thinking. It seems that one of the prototypical characteristics of higher cognition is representing abstract relations, simulating events, and reasoning about the consequences that have yet to happen. If we look to what causal reasoning involves, mental simulation is often required to play out various scenarios and to imagine hypothetical consequences. Is this capacity associated with T1 or T2? Take our ability to follow complex story lines in soap operas (Hagmayer & Osman, in press). We are able to mentally simulate multiple complex events and infer what would happen next with a great deal of ease. In fact, soap operas work precisely because of this capability, otherwise we would not be able to follow story lines or even be engaged in what goes on. In these and other social contexts, evidence shows that causal knowledge is recruited to help mentally simulate the various outcomes of different hypothetical situations. Moreover, high level mental simulation can be achieved with very little processing cost. So, one question that this work raises is how the ability to simulate over multiple hypothetical scenarios is reconciled within E&S's DPT, given that mental simulation is definitively T2 processing, and immunity from loading of WMC is definitively T1 processing.

Finally, turning to decision making in the real world, one program of research born out of qualitative distinctions between T1 and T2 has now moved on from making dichotomies (Osman, 2010a, 2010b). For instance, attempts to find qualitative dissociations between dynamic decision making based on WMC capacity and automaticity have not stood the test of time and have since been abandoned. The research agenda and research questions are focused on issues that have moved away from dissociations, because the complexity of the cognition that is being studied cannot be boxed in to such a narrow conception. Perhaps the same trend may apply to other higher cognitive functions (e.g., reasoning, judgment, problem solving). It is very likely that future research will show that these higher cognitive functions are each multifaceted and are built on other multifaceted processes. If so, is it appropriate to carve up higher cognitive functions into such a simple framework (DPTs)? Especially when the distinctions between T1 and T2 are based on what are actually quantitative rather than qualitative differences?

References

De Neys, W. (2006a). Automatic-heuristic and executive-analytic processing in reasoning: Chronometric and dual task considerations. *Quarterly Journal of Experimental Psychology*, *59*, 1070–1100.

De Neys, W. (2006b). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433.

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, *7*, 28–38.

De Neys, W., & Dieussaert, K. (2005, July). *Individual differences in rational thinking time*. Paper presented at the Annual Meeting of the Cognitive Science Society, Stresa, Italy.

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

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences*, *7*, 454–459.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, England: Psychology Press.

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

Hagmayer, Y., & Osman, M. (in press). From colliding billiard calls to colluding desperate housewives: Causal Bayes nets as rational models of everyday causal reasoning. *Synthese*.

Hommel, B. (2000). The prepared reflex: Automaticity and control in stimulus-response translation. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 247–273). Cambridge, MA: MIT Press

Klauer, K., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, *107*, 852–884.

Oaksford, M., Morris, F., Grainger, B., & Williams, J. M. G. (1996). Mood, reasoning, and central executive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 476–492.

Osman, M. (2004). An evaluation of dual process theories of reasoning. *Psychonomic Bulletin & Review*, *11*, 998–1010.

Osman, M. (2007). Can tutoring improve performance on a reasoning task under deadline conditions? *Memory & Cognition*, *35*, 342–351.

Osman, M. (2010a). *Controlling uncertainty: Learning and decision making in complex worlds*. Chichester, England: Wiley-Blackwell.

Osman, M. (2010b). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin*, *136*, 65–86.

Posner, M. (1978). *Chronometric exploration of mind*. Hillsdale, NJ: Erlbaum.

Roberts, M. J., & Newton, E. J. (2001). Inspection times, the change task, and the rapid-response selection task. *Quarterly Journal of Experimental Psychology*, *54A*, 1031–1048.

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

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago, IL: University of Chicago Press.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, *23*, 645–726.

Wilkins, M. C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology*, *102*, 1–83.