The Handbook of Rationality

Edited by: Markus Knauff, Wolfgang Spohn

Citation:

The Handbook of Rationality
Edited by: Markus Knauff, Wolfgang Spohn
DOI: 10.7551/mitpress/11252.001.0001
ISBN (electronic): 9780262366175
Publisher: The MIT Press
Published: 2021

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.

The MIT Press

2.5 Dual-Process Theories of Deductive Reasoning

Karl Christoph Klauer

Summary

In dual-process models of thinking and reasoning, two kinds of processes are distinguished. Defining and correlated characteristics of these processes as well as their different functions are discussed along with the tasks and phenomena typically addressed by dual-process models. The relationships between the types of processes and normative, instrumental, and epistemic rationality are considered. A selective review of specific dual-process models illustrates the range of dual-process models that have been proposed. Evidence for and against dual-process models is summarized and evaluated.

1. Two Types of Thinking and Reasoning

Dual-process models of thinking and reasoning postulate that two kinds of processes are involved in reasoning. These two kinds of processes differ qualitatively and interact or compete in generating responses to reasoning problems. Dual-process models of thinking and reasoning are part of a larger family of such models in areas as diverse as social cognition, learning, and moral decision making. The guiding idea is that one kind of thinking, Type 1, delivers intuitive responses, gut feelings, or fast heuristic response proposals, whereas the other, Type 2, is responsible for slow but careful and systematic deliberations. Type 1 processes have been described as heuristic, context dependent, fast, and automatic; Type 2 processes as analytic, context independent, slow, and controlled.

In processing a reasoning problem, Type 1 processes are responsible for encoding the premises, for directing attention to relevant aspects in a bottom-up stimulus-driven manner, and for proposing an intuitive response. Furthermore, Type 1 processes have been argued to replace and add premises to reasoning problems through pragmatic implicatures and background knowledge. Finally, a Type 1 response can also function as a fallback response if Type 2 processes fail to propose a response.

Type 2 processes are involved in processing the task instructions. They assemble a task-set as an ad hoc constellation of processes coordinated to satisfy the task demands. Furthermore, Type 2 processes constantly monitor performance according to the implemented task-set and higher-level goals. They are responsible for strategic allocation of attentional resources and for processing that requires working-memory resources and mental simulation. They are activated when, for example, Type 1 processes do not suggest a response, suggest conflicting responses, or register surprise or a threat (e.g., Kahneman, 2011).

Within this framework, several dual-process models have been proposed that differ, among other issues, on how precisely Type 1 and Type 2 processes are characterized. Dual-process models have to explain as a minimum what distinguishes the two kinds of reasoning and how they interact in generating responses to reasoning problems.

Much debate has focused on which aspects distinguish Type 1 from Type 2 processes; table 2.5.1 provides a less than complete list of distinctions compiled in large part from a similar table in Evans and Stanovich (2013a). Furthermore, there is debate on how the two types of processes interact. One kind of interaction is that Type 1 processes shape the mental representations on which Type 2 processes operate (e.g., Evans, 2009; Thompson, 2013) by drawing attention to certain features of the problem at the expense of others and by adding premises from background knowledge. Type 1 processes also often generate a response proposal for the reasoning problem. In a default-interventionist model, this proposal serves as a default response that is acted upon unless Type 2 processes intervene (Evans, 2007). In contrast, in the parallel-competitive model, both kinds of processes operate in parallel, producing response proposals that compete for control of the eventual response (Evans, 2007).

Table 2.5.1
Some features and dimensions with respect to which Type 1 and Type 2 processes have been argued to differ

Type 1 process	Type 2 process
Fast	Slow
Automatic	Controlled
Nonconscious	Conscious
Unintentional	Intentional
Associative	Reflective
Parallel	Serial
Heuristic	Analytic
Experiential	Rational
Does not require working memory	Requires working memory

2. Stanovich's Tripartite Model

It is helpful to add a number of conceptual distinctions to this framework. In doing so, I will draw heavily on Stanovich's (2009) treatment. Stanovich (2009) defines the entirety of Type 1 processes as the set of autonomous systems in the brain that operate autonomously in response to specific triggering stimuli. These comprise many different and heterogeneous kinds of processes such as

- preattentive processes that supply content for focal attention and thereby shape the initial mental representation of the problem (Evans, 2009),
- domain-specific processing modules as discussed by evolutionary psychologists (e.g., Cosmides, 1989),
- domain-general processes of implicit learning and conditioning, and
- inference, decision-making, and classification rules (Mercier & Sperber, 2009) that are innate or have been practiced to automaticity (Anderson, 1983; Logan, 1985).

Type 2 processes, on the other hand, fall into two classes, termed the "algorithmic" and the "reflective" mind. The reflective level comprises control states that regulate processing according to higher-level goals. Processes at the reflective level initiate and guide processes at the algorithmic level. The algorithmic level implements processes and strategies such as those required in solving the tasks of a test of fluid intelligence, and it sustains hypothetical thought and mental simulations. The algorithmic processes can also inhibit and override the output of Type 1 processes. The reflective processes, however, initiate such overriding as well as mental simulations.

Stanovich (2009) distinguishes between the algorithmic and the reflective level in part on theoretical grounds: they serve different functions in reasoning. But the distinction is also empirically motivated by individual-differences research on reasoning performance in different tasks (see Stanovich, 2009).

Finally, Stanovich's (2009) framework allows him to distinguish between different kinds of reasoning biases:

- First of all, Type 2 processes may be not engaged. This can come about where a biased Type 1 output is so compelling (for example, associated with high feelings of rightness; Thompson, 2009) that Type 2 processes are not invoked or where Type 2 processes are not available due to lack of capacity or effort.
- Type 2 processes may be engaged but fail to override a biased Type 1 response. Lack of effort is not the only possible cause of this outcome. Alternatively, the necessary Type 2 rules and algorithms may not have been learned and may therefore not be available, a situation that Stanovich (2009) called "mindware gaps."
- Relatedly, the available "mindware" may be faulty or itself biased, as in Stanovich's (2009) category of "biases due to contaminated mindware."
- Type 2 processes may operate on a mental representation that does not permit one to find the correct solution. For example, in Wason's (1966) selection task, shown in figure 2.5.1, preattentive Type 1 processes may draw attention to the cards whose visible sides show the elements explicitly mentioned in the rule ("A" and "3" in figure 2.5.1), which do not include one of the cards that needs to be selected ("1" in figure 2.5.1). Any reasoning focusing on only these cards would probably be seen as an instance of rationalization of an initial response (Evans, 1996; Evans & Wason, 1976). Stanovich (2009) terms this kind of Type 2 reasoning "serial associative cognition with a focal bias."

3. Type of Process and Rationality

3.1 Instrumental, Normative, and Epistemic Rationality

What is the relationship of the different processes to human rationality? Consider instrumental rationality, defined as choosing suitable means for one's ends. Type 1 processes, if defined as autonomous and triggered by specific cues, have no clear relationship to instrumental

The (reduced) base-rate task
(after Pennycook, Fugelsang, & Koehler, 2015)

General Instructions:
In a big research project a large number of studies were carried out where short personality descriptions of the participants were made. In every study there were participants from two population groups (e.g., carpenters and policemen). In each study one participant was drawn at random from the sample. You'll get to see a personality trait for this randomly chosen participant. You'll also get information about the composition of the population groups tested in the study in question. You'll be asked to indicate to which population group the participant most likely belongs.

One Trial:
This study contains: politicians and nannys.
Person 'A' is dishonest.
There are 999 nannys/5 politicians.
Is person 'A' more likely to be a) a politician or b) a nanny?

The belief-bias task
(after Handley, Newstead, & Trippas, 2011)

General Instructions:
You must assume each premise is true (even if in reality it is not true) and respond with the answer which logically follows from the statements presented.

For example: If you finish a drink, then the drink will be full.

Suppose that you finish your drink.
Will your drink be empty?

The logic-based answer to this problem is NO, because you must assume that if a drink is finished, then it will be full; therefore, the glass will not be empty.

One Trial:
If a child is crying, then he is happy.
Suppose a child is crying.
Is the child happy?

The Wason selection task
(after Klauer, Stahl, & Erdfelder, 2007)

Instructions:
Below you see a number of cards from a set of cards. Each card in the set has a capital letter on one side and a digit on the other. Naturally, only one side is visible in each case. For the set of cards, a rule has been stated. It is: "If there is an A on the letter side of the card, then there is a 3 on the number side." You must decide which card(s) displayed would have to be turned over in order to test the truth or falsity of the rule.

A **B** **3** **1**

Figure 2.5.1
Three tasks frequently used in research on dual-process models.

rationality as they are not directly informed by one's goals. The same is true for normative rationality, defined as the adherence to given norms. There are, however, a couple of considerations that qualify these strong statements. Type 1 processes, it is now widely held, can also deliver normatively correct responses if they are well attuned to environmental regularities, and they can be sensitive to abstract structural features of a task. For example, the modus ponens rule describes the inference from a major premise of the form "If p then q" and a minor premise "p" to the conclusion "q." Modus ponens may be so well practiced that it can often be triggered habitually and fast in reading and comprehending the premises (e.g., Lea, 1995; Lea, O'Brien, Fisch, Noveck, & Braine, 1990; Rader & Sloutsky, 2002).

Relatedly, Type 1 processes are frequently argued to allow us to approach major goals that we hold (e.g., Evans, 2014), linking them to instrumental rationality. This is sometimes associated with dual-systems views or two-minds theories in which Type 1 and Type 2 processes live in different systems, linked to different evolutionary stages (e.g., Evans, 2014; Mercier & Sperber, 2017; Stanovich, 2004).

For example, integrating one's beliefs about the likelihood of possible outcomes of acts and one's evaluations of these outcomes is seen as a hallmark of instrumental rationality in decision theory (Buchak, 2013). At the same time, relying on one's beliefs underlies the generation of belief bias (Wilkins, 1928) in deductive reasoning, which has sometimes been construed as reflecting the operation of a Type 1 process. Similarly, Gigerenzer and Todd (1999) have argued (a) that human beings have at their disposal a toolbox of fast and frugal heuristics, some of which would probably qualify as Type 1 processes for some authors, and (b) that these often mimic optimal strategies for achieving current goals by capitalizing on regularities in benign environments. And so, instrumental rationality alone is unlikely to be a good feature for cleanly distinguishing Type 1 from Type 2 processes. Similar points can be made regarding normative rationality.

Epistemic rationality regards the internal coherence and consistency of the beliefs that we hold and the support that we seek for them in terms of external evidence. In his two-minds theory, Evans (2014) argues that epistemic rationality is something pursued only by what he calls the "new" mind and Type 2 processes operating on explicit knowledge. In contrast, both the "old" mind, relying on Type 1 processes, and the "new" mind work in concert to achieve instrumental rationality. The next section explores the idea that justifications as an element of epistemic rationality might reliably distinguish between Type 1 and Type 2 processes.

3.2 Justification as a Qualitative Distinction between Type 1 and Type 2 Processes?

Consider an analogy to Plato's definition of knowledge as *justified* true belief. Type 1 processes can deliver correct responses (true belief), satisfying the requirements of instrumental rationality and sometimes normative rationality. But they do so for the wrong reasons (i.e., unjustified) inasmuch as they capitalize on environmental features that happen to covary with the normative or rational response most of the time. This can be revealed, for example, in experimental situations in which the auspicious environmental features are intentionally removed and in which responses then depart from the normative prescriptions or from those of instrumental rationality. Kahneman and Tversky's heuristics-and-biases program (see chapter 2.4 by Fiedler, Prager, & McCaughey, this handbook) has been characterized as having achieved just this (Kahneman, 2011). For example, one's beliefs will often provide an adequate guideline when judging the validity of the conclusion of a logical argument, but they may be misleading when the task is to disregard one's background knowledge and to judge the validity solely based on whether the proposition follows logically from the premises of the argument, taking these as true (see figure 2.5.1 for an example).

And thus, pursuing the above analogy a bit further, it may not be enough to focus on the correctness or reasonableness of the outcomes of the reasoning process. Instead, the outcome should also come about for the right reasons. One way to ensure this is to require that reasoners should be aware of reasons or justifications for the particular reasoning steps they engage in. This requirement to be able to justify one's responses as they are generated (rather than post hoc) captures nicely the distinction between Type 1 processes, characterized as intuitions being based on gut feelings or reactions, and deliberate Type 2 processes, based on reflection. In a similar vein, Mercier and Sperber (2009) suggest distinguishing between intuitive and reflective inferences. Reflective inferences involve the representation of reasons for the inferred conclusions (see also Mercier & Sperber, 2017). Of course, conclusions that come with justifications need not be normatively or instrumentally correct, for reasons already discussed above, but this does not diminish their potential as a demarcation criterion.

4. Examples of Dual-Process Models and Related Models in the Reasoning Literature

4.1 Evans and Stanovich's Framework

Let us consider a number of dual-process models in the reasoning domain to illustrate these concepts further. Evans and Stanovich (2013a) have provided a framework for dual-process models in reasoning and decision making in which the two types of reasoning are distinguished in terms of two defining features: Type 2 processes load working memory, whereas Type 1 processes do not require working-memory resources. In addition, another hallmark of Type 2 processing is that it sustains hypothetical thought or cognitive decoupling (Stanovich, 2009), "the ability to distinguish supposition from belief and to aid rational choices by running thought experiments" (Evans & Stanovich, 2013a, p. 236). In particular, working memory and cognitive decoupling are involved when Type 1 processes and their outputs are to be inhibited.

The interaction between the two types of processing is described in terms of a default-interventionist model: Type 1 processes deliver a default intuition that can be overridden by Type 2 processes. Whether and when this occurs is a function of motivation, the availability of working-memory resources, and individual differences in thinking dispositions (Evans & Stanovich, 2013b). Type 2 processing can be algorithmic, proposing an alternative response to the default intuitive one or affirming it, and reflective, embodying goal-driven higher-level control responsible for the decision to override Type 1 processing. Evans and Stanovich (2013b) see this framework as a paradigm or metatheory that can inspire task-level theories.

4.2 A Single-Function Dual-Process Model

Oaksford and Chater (2010, 2012) propose what they called a "single-function dual-process model" targeted at reasoning with conditionals. Type 1 processes in this model are the processes of spreading activation in a constraint-satisfaction network identified with long-term memory. Type 2 processes operate on both long-term memory and working memory and are responsible for strategic queries of long-term memory, for representing the outputs of these queries in working memory, and for operations on these working-memory contents.

Upon encountering a conditional sentence of the form “If *p* then *q*,” a part of long-term memory is activated, including the nodes for the antecedent (*p*) and the consequent (*q*) as well as other representations strongly linked with them. Given a minor premise, such as the antecedent in a modus ponens inference, reasoners will clamp on the node representing the minor premise and monitor the flow of activation in the activated portion of long-term memory until an equilibrium is reached. The activation of the node representing the consequent is then read off and the result of the Ramsey-style test (Bennett, 2003) stored in working memory. This representation in working memory is actually conceived of as similar to the mental-model representations in Johnson-Laird’s mental model theory (see chapter 2.3 by Johnson-Laird, this handbook).

Strategic queries of long-term memory allow the constraint-satisfaction network to compute activation levels for hypothetical and counterfactual premises. Type 2 processes are also responsible for undoing changes in long-term memory due to a previous query, and they can have a decoupling function in screening the representations in working memory from the intrusion of contextualized long-term memory information. This can result in processing that mimics logical thought as a limiting case of the function primarily computed, namely conditional probabilities. Nevertheless, the two kinds of processes serve a single function: to assess probabilities.

Type 1 and Type 2 processes thus differ on several of the features listed in table 2.5.1, most conspicuously in terms of whether working memory is required. Furthermore, although Oaksford and Chater (2010) do not explicitly say so, processes of spreading activation in long-term memory would presumably be nonconscious and, once triggered by stimuli or a Type 2 query, run to completion automatically. In contrast, Type 2 processes would probably be seen as intentional and conscious attempts to probe long-term memory and to represent and integrate mental-model-like representations in working memory. The interaction of the two processes is twofold: Type 1 processes determine which parts of long-term memory are activated, shaping the problem representation. Furthermore, the interaction can take the default-interventionist form in that Oaksford and Chater (2010) assume that a default query of long-term memory can occur stimulus driven and provide a default assessment of the relevant probabilities that can however be modified by Type 2 queries of long-term memory to the point where the default assessment is completely inhibited.

4.3 A Dual-Strategy Model

Verschueren, Schaeken, and d’Ydewalle (2005a, 2005b) have proposed a dual-strategy model of conditional reasoning that was further developed by Markovits and colleagues (e.g., Markovits, Brunet, Thompson, & Brisson, 2013; see also chapter 15.1 by Markovits, this handbook). The idea is that people can use a combination of statistical and counterexample forms of reasoning. Markovits et al. (2013) align the statistical strategy with the assessment of probabilities similar to the stimulus-driven long-term memory queries considered by Oaksford and Chater (2010, 2012). The statistical strategy is to accept conclusions that are assessed as probable (somewhat like in accepting believable conclusions in belief-bias problems, discussed in more detail below). It is associated with little cognitive cost, operates fast, and delivers a default response. Given time and cognitive capacity, the statistical strategy can be replaced by a counterexample strategy in which a conclusion is rejected if a counterexample can be found. Counterexamples can be implied by the given information, but as in Oaksford and Chater’s (2010, 2012) single-function dual-process model, counterexamples can also be generated via a strategic search of relevant background knowledge. The counterexample strategy is slower, and it comes with a higher cognitive cost. One difference to Oaksford and Chater’s approach is that the two kinds of processes compute different functions, the former being sensitive to probabilistic information, the latter being driven by the existence or absence of at least one counterexample. The approach is thus a dual-function dual-process one.

4.4 A Dual-Source Model

Consider finally the dual-source model of probabilistic conditional reasoning by Klauer, Beller, and Hütter (2010; see also Singmann, Klauer, & Beller, 2016). The model distinguishes between two processes in terms of whether they derive response proposals based on background knowledge or based on the form of the argument. The model is not committed to further possible distinctions between the two types of processes. Instead, it provides a methodology and experimental paradigm to disentangle the contributions of both kinds of processes, making them and their features amenable to empirical study.

5. Evidence in Favor of and Against Dual-Process Models

5.1 Working-Memory Load, Speeded Responses, and Patterns of Interference

Evidence in favor of dual-process models of reasoning has focused on tasks such as those shown in figure 2.5.1. In a classical belief-bias task, participants are to judge the validity of valid and invalid syllogisms with believable and unbelievable conclusions. Belief bias is the finding that participants are more likely to accept as logically valid syllogisms with believable conclusions than syllogisms with unbelievable conclusions. As already mentioned, the assessment of believability has been argued to be based on a Type 1 process that is fast and requires little working-memory involvement, whereas logic-based responses require slower and working-memory-dependent Type 2 processes. In line with these ideas, both time pressure (Evans & Curtis-Holmes, 2005) and working-memory load (De Neys, 2006) decreased participants' ability to discriminate between valid and invalid syllogisms, whereas the effect of believability was increased (Evans & Curtis-Holmes, 2005) or not affected (De Neys, 2006).

On the other hand, Handley, Newstead, and Trippas (2011) also asked a group of participants to judge the believability of the conclusion and found (a) that under some conditions, judgments of believability took more time than judgments of logical validity and (b) that judgments of believability were interfered with by the validity of the syllogism in question: conclusions of valid syllogisms were judged more believable than conclusions of invalid syllogisms. This suggests a more symmetric role for the belief-based and the logic-based process than the studies just cited.

Let us take a closer look at the actual problems employed. Evans and Curtis-Holmes (2005) and De Neys (2006) employed problems with conclusions that were either believable or unbelievable. For example, a believable conclusion was "Some healthy people are not astronauts"; a (relatively more) unbelievable conclusion was "Some astronauts are not healthy people." Judging the believability of these conclusions is probably driven by a spontaneous, stimulus-triggered memory query of the kind described by Oaksford and Chater's (2010) dual-process models. In contrast, Handley et al. (2011, experiments 1 to 3) employed the belief-bias problems exemplified in figure 2.5.1 in which the conclusion (e.g., "The child is happy") is neither believable or unbelievable. It only acquires a value on the believability dimension once the minor premise ("The child is crying") is integrated with relevant conditional background knowledge (such as "If a child is crying, then he or she is not happy"). Note that assessing the believability of the conclusion thereby requires almost the same processing steps as assessing its validity, the only difference being that the minor premise needs to be integrated with the explicitly stated conditional rule ("If a child is crying, then he is happy") in checking validity, whereas it needs to be integrated with conditional background knowledge in assessing believability. In short, assessing validity and assessing believability appear to have been equated in terms of processing demands, questioning their alignment with different types of processing. Note also that this issue does not afflict Handley et al.'s experiments 4 and 5, which are, however, subject to a similar alternative account in terms of aftereffects of switching between assessing believability and validity from trial to trial (Handley et al., 2011, p. 40). Switching is known to inflate existing, and sometimes newly create, crosstalk between the two alternating tasks (Allport, Styles, & Hsieh, 1994).

Similar ideas have been further developed in Chun and Kruglanski's (2006; see also Kruglanski & Gigerenzer, 2011) work on the base-rate task. In base-rate problems, participants are to decide which of two professional categories a person is more likely a member of. Two kinds of information are provided: first, participants are told that the person in question is randomly drawn from a sample of persons with skewed base rates for the two categories. For example, there might be 800 lawyers and 200 engineers. The second source of information is a description of traits and/or habits of the person such as that the person is not married, is introverted, and likes to spend his free time reading science fiction and writing computer programs. The second source is thus a description that is stereotypical of members of one of the two categories, engineers in the present example. Relying on the stereotypical information is usually considered heuristic, processing the base-rate information can be analytical. A typical finding is that the base-rate information is neglected and only the stereotypical information used (Kahnemann & Tversky, 1973).

Chun and Kruglanski (2006) argue, however, that the decisive difference between the two kinds of information is not related to the different contents (base rates versus stereotypes) but to how easy it is to process them. In general, equating how easy it is to process the two kinds of information should make their impact on judgments in the base-rate task more symmetrical, as Chun and Kruglanski (2006) demonstrate across several studies.

We argued that this partly explains Handley et al.'s (2011) results: the task demands of some of their belief-bias problems equated processing ease for judging validity and judging believability. Indeed, when more difficult logical problems are used, the effects of validity on judgments of belief decrease, as shown by Trippas, Thompson, and Handley (2017), although these studies are also afflicted by the abovementioned switching confound that is known to inflate crosstalk between the two alternating tasks.

A similar issue may be in place for many recent studies using the base-rate task (e.g., Bago & De Neys, 2017; De Neys, 2012, 2014; De Neys & Glumicic, 2008; Newman, Gibb, & Thompson, 2017; Pennycook, Fugelsang, & Koehler, 2015; Pennycook & Thompson, 2012; Pennycook, Trippas, Handley, & Thompson, 2014). In the "reduced" version shown in figure 2.5.1, the issue is especially clear: base-rate information and stereotypical information are both presented in one sentence. Both can suggest responses via simple rules or heuristics—to pick the larger category for the base-rate information and to pick the matching category for the stereotypical information. Again, ease of processing appears to be closely comparable for the two heuristics, questioning their alignment with different kinds of processing.

For example, Pennycook et al. (2015) erect a three-stage dual-process model on complex patterns of interferences in response times and response rates in the reduced base-rate task. But as pointed out by Krajbich, Bartling, Hare, and Fehr (2015), these patterns of interference can also come about through simpler one-process models in which the strength of preference (operationalized in terms of relative choice frequency) for one of the responses (the stereotypical response versus the base-rate response) and response time are related so that the preferred response is generally given faster the stronger the preference for it. Pennycook, Fugelsang, Koehler, and Thompson (2016) acknowledge this argument and point to aspects of their data pattern that may still be difficult to account for by this alternative account.

A number of these studies also used a two-response paradigm (e.g., Bago & De Neys, 2017) in which participants give a first response as quickly as possible followed by a second response without time pressure. One major finding is that most participants stick to their first responses, so that changes from first to second response are rare. This has been argued to question a number of dual-process models: if the first response relies on Type 1 processes, whereas Type 2 processes are involved in the second response, changes should regularly occur when both kinds of processes suggest different responses. One clue to what is happening here may be that participants appear to be very consistent in their responses to the problems (e.g., De Neys, 2017a, p. 56): across problems, they either rely on base-rate (or in other cases logical) information or on stereotypical information, suggesting that participants decide and perhaps justify to themselves early which of these two sources of information to rely on. They then stick to this deliberate decision rather than reevaluate it for each problem anew. If both sources of information are easy to process, both strategies are likely already available for the first response.

In fact, when problems are used in which logic and believability, or base-rates and stereotypical information, are arguably less strongly equated in processing difficulty, changes are more frequent (Thompson, Prowse Turner, & Pennycook, 2011). In addition, in Thompson et al.'s (2011) studies, feeling-of-rightness (FOR) judgments for the first responses significantly predicted how often changes occur and how long it takes to produce the second response after the first. In a default-interventionist dual-process framework, this is consistent with the idea that FOR determines the likelihood and intensity of the engagement of Type 2 processes.

5.2 Correlational Evidence

Another source of evidence is individual-differences research. Stanovich and West (2000; see also chapter 15.2 by Stanovich, Toplak, & West, this handbook) in particular have investigated correlates of task performance in many tasks in which the modal response deviates from the one considered normative. In many of these tasks, measures of intelligence correlated positively with normative correctness of the response and negatively with giving the modal response. Importantly, predictions can be made for where to expect such correlations (e.g., when Type 2 processing is substantially involved in a version of the task) and where not (e.g., when Type 1 processing usually generates the correct response without conflict with Type 2 processing)—predictions that have been confirmed in several instances (Evans & Stanovich, 2013a, 2013b).

6. Criticisms of Dual-Process Models

6.1 Distinguishing Type 1 and Type 2 Processes: The Dichotomous and the Gradual View

Criticisms of dual-source models have addressed many aspects of dual-process models (Keren, 2013; Keren & Schul, 2009; Kruglanski & Gigerenzer, 2011). A number of problems identified by Keren and Schul (2009) and Keren (2013) arise from the fact that dual-process models have historically characterized the two types of processes in terms of arrays of attributes (see table 2.5.1) that are not perfectly correlated, an approach explicitly eschewed in recent theoretical expositions of dual-process theory (e.g., Evans & Stanovich, 2013a).

This reminds one of earlier debates about the definition of automaticity (for a review, see Moors & De Houwer, 2006): some definitions understood the contrast between automatic and controlled processes as a dichotomy (Shiffrin & Schneider, 1977) with different clusters of features characterizing automatic and controlled processes, respectively, much like exemplified in table 2.5.1 for Type 1 and Type 2 processes. Other definitions considered a continuum of attentional requirements with automatic processes at one end and nonautomatic processes at the other (e.g., Anderson, 1983). The dichotomous view was challenged by studies finding lack of correlation among central features ascribed to automatic processes, such as that they are efficient, unintentional, uncontrollable, and unconscious (Bargh, 1992). Such challenges led to a decompositional view according to which different automaticity features are to be studied separately and independently, suggesting in the end the deconstruction of the concept of automaticity.

Gradual views, on the other hand, maintain the concept and align it with a transition from algorithmic computation to single-step memory retrieval (Logan, 1988) or algorithmic strengthening in skill acquisition (Anderson, 1983). Logan (1985) argued that the different automaticity features have different time courses of change with practice, and thus lack of co-occurrence among the features is no surprise. The gradual view would only be challenged if different features were found to develop in different directions as practice increases. Whereas the gradual view makes it difficult to draw a line distinguishing automatic from nonautomatic processes, another response was to raise one dichotomous feature to a definitional status and acknowledge that other features might, but need not, co-occur with it (Bargh, 1992).

Similar arguments and moves can be found in debates on dual-process models. For example, some, although not all, of the features in table 2.5.1 are inherently gradual and define processing continua rather than qualitative distinctions (Keren, 2013; Keren & Schul, 2009). Involvement of working-memory resources, a feature highlighted by Evans and Stanovich (2013a), clearly comes in degrees, and it is notoriously difficult to distinguish between small and zero involvement. In a similar vein, Kruglanski and Gigerenzer (2011; see also Chun & Kruglanski, 2006) acknowledge that different processes differ on many qualitative aspects such as the contents that they operate on but that graded differences in processing ease and in perceived ecological validity may be responsible for relatively greater impacts on the ultimate response of some processes than others. Cognitive decoupling and mental simulations, another distinguishing feature highlighted by Evans and Stanovich (2013a), as well as the requirement to provide reasons (Mercier & Sperber, 2009), may be more difficult to construe as gradual distinctions, depending upon how exactly these features are defined and operationalized, and they may serve as defining features if a dichotomous view is preferred.

6.2 The Relative View

Alternatively, the conceptualization of dual-process models could also be based on a relative view as proposed by Moors and De Houwer (2006) for the distinction between automatic and controlled processes. This is a refinement of the gradual view in which a standard of comparison is introduced. In terms of dual-process models, a process would be called Type 1 relative to a standard of comparison appropriate to a given task context, whereas in another context, the same process might be classified as Type 2 relative to a standard of comparison that is prevalent in that context. For example, in the context of the Stroop task, a word-reading and a color-naming process are involved. When the task is to name the color in which a word is written, uninstructed, spontaneous reading of the word interferes with naming the word's color if the word itself denotes a different color. In contrast, if the task is to read the word, there is no analogous interference from the word's color in conflict trials (MacLeod, 1991). To account for this asymmetry, word reading has been considered relatively more automatic (Type 1) and color naming relatively more controlled (Type 2; Lindsay & Jacoby, 1994). In the tasks studied in the thinking-and-reasoning literature, both processes would, however, probably be considered Type 1 processes relative to the resource-depleting reasoning algorithms demanded by many reasoning tasks. Similarly, assessing believability in belief-bias tasks may normally tax cognitive resources so much less than assessing logical validity that the relative view would designate the former process Type 1 and the latter Type 2. If and when assessing believability is, however, made more difficult, and assessing logical validity easier, thereby equating both in terms of processing demands (as in Handley et al., 2011, experiments 1 to 3), the gradual view would deny such a designation in this context.

7. Conclusion

Keren (2013) argues that single-process accounts may be offered for a number of dual-process phenomena (see also Chun & Kruglanski, 2006; Krajbich et al., 2015; Kruglanski, 2013). Proponents of dual-process models need to defend their theories against this criticism, but the criticism is targeted at specific instances of dual-process models for specific tasks. At a more abstract level, dual-process models have proven to provide an immensely fruitful framework for integrating a diversity of related findings. They serve as a template for devising specific accounts of tasks of interest, providing a rich toolbox of concepts and process assumptions (Evans & Stanovich, 2013b). The framework has thereby generated an enormous amount of research producing new tasks and revealing new phenomena that might not have been uncovered without it, many of them summarized in the recent book by De Neys (2017b). There is little doubt that the core ideas and concepts of dual-process models will continue to be influential and to open up new and exciting avenues of research.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

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

Bago, B., & De Neys, W. (2017). Fast logic? Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic effects. *American Journal of Psychology*, *105*, 181–199.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford, England: Clarendon Press.

Buchak, L. (2013). *Risk and rationality*. Oxford, England: Oxford University Press.

Chun, W. Y., & Kruglanski, A. W. (2006). The role of task demands and processing resources in the use of base-rate and individuating information. *Journal of Personality and Social Psychology*, *91*, 205–217.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187–276.

De Neys, W. (2006). Dual processing in reasoning. *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. (2014). Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning*, *20*, 169–187.

De Neys, W. (2017a). Bias, conflict, and fast logic: Towards a hybrid dual process future? In W. De Neys (Ed.), *Dual process theory 2.0* (pp. 47–65). Abingdon, England: Routledge.

De Neys, W. (Ed.). (2017b). *Dual process theory 2.0*. Abingdon, England: Routledge.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248–1299.

Evans, J. St. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology*, *87*, 223–240.

Evans, J. St. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, *13*(4), 321–339.

Evans, J. St. B. T. (2009). How many dual-process theories do we need? One, two, or many? In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33–54). New York, NY: Oxford University Press.

Evans, J. St. B. T. (2014). Two minds rationality. *Thinking & Reasoning*, *20*(2), 129–146.

Evans, J. St. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, *11*, 382–389.

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

Evans, J. St. B. T., & Stanovich, K. E. (2013b). Theory and metatheory in the study of dual processing: Reply to comments. *Perspectives on Psychological Science*, *8*, 263–271.

Evans, J. St. B. T., & Wason, P. C. (1976). Rationalization in a reasoning task. *British Journal of Psychology*, *67*, 479–486.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, & ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 3–34). New York, NY: Oxford University Press.

Handley, S. J., Newstead, S. E., & Trippas, D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 28–43.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Keren, G. (2013). A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*, 257–262.

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science*, *4*, 533–550.

Klauer, K. C., Beller, S., & Hütter, M. (2010). Conditional reasoning in context: A dual-source model of probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 298–323.

Klauer, K. C., Stahl, C., & Erdfelder, E. (2007). The abstract selection task: New data and an almost comprehensive model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 680–703.

Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, *6*.

Kruglanski, A. W. (2013). Only one? The default interventionist perspective as a unimodel—Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*, 242–247.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*, 97–109.

Lea, R. B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1469–1482.

Lea, R. B., O'Brien, D. P., Fisch, S. M., Noveck, I. A., & Braine, M. D. S. (1990). Predicting propositional logic inferences in text comprehension. *Journal of Memory and Language*, *29*, 361–387.

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 219–234.

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *39*, 367–386.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.

Markovits, H., Brunet, M.-L., Thompson, V., & Brisson, J. (2013). Direct evidence for a dual process model of deductive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1213–1222.

Mercier, H., & Sperber, D. (2009). Intuitive and reflective inferences. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 149–170). New York, NY: Oxford University Press.

Mercier, H., & Sperber, D. (2017). *The enigma of reason: A new theory of human understanding.* Cambridge, MA: Harvard University Press.

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

Newman, I. R., Gibb, M., & Thompson, V. A. (2017). Rule-based reasoning is fast and belief-based reasoning can be slow: Challenging current explanations of belief-bias and base-rate neglect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1154–1170.

Oaksford, M., & Chater, N. (2010). Conditionals and constraint satisfaction: Reconciling mental models and the probabilistic approach? In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 309–334). Oxford, England: Oxford University Press.

Oaksford, M., & Chater, N. (2012). Dual processes, probabilities, and cognitive architecture. *Mind & Society*, *11*, 15–26.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72.

Pennycook, G., Fugelsang, J. A., Koehler, D. J., & Thompson, V. A. (2016). Commentary: Rethinking fast and slow based on a critique of reaction-time reverse inference. *Frontiers in Psychology*, *7*, Article 1174.

Pennycook, G., & Thompson, V. A. (2012). Reasoning with base rates is routine, relatively effortless, and context dependent. *Psychonomic Bulletin & Review*, *19*, 528–534.

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 544–554.

Rader, A. W., & Sloutsky, V. M. (2002). Processing of logically valid and logically invalid conditional inferences in discourse comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 59–68.

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.

Singmann, H., Klauer, K. C., & Beller, S. (2016). Probabilistic conditional reasoning: Disentangling form and content with the dual-source model. *Cognitive Psychology*, *88*, 61–87.

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

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55–88). New York, NY: Oxford University Press.

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

Thompson, V. A. (2009). Dual-process theories: A metacognitive perspective. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 171–196). New York, NY: Oxford University Press.

Thompson, V. A. (2013). Why it matters: The implications of autonomous processes for dual-process theories—Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, *8*, 253–256.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140.

Trippas, D., Thompson, V. A., & Handley, S. J. (2017). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory & Cognition*, *45*(4), 539–552.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005a). A dual-process specification of causal conditional reasoning. *Thinking & Reasoning*, *11*(3), 239–278.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005b). Everyday conditional reasoning: A working memory–dependent tradeoff between counterexample and likelihood use. *Memory & Cognition*, *33*(1), 107–119.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (Vol. 1, pp. 135–151). Harmondsworth, England: Penguin.

Wilkins, M. C. (1928). *The effect of changed material on ability to do formal syllogistic reasoning* (Archives of Psychology No. 102). New York, NY: Woodworth.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Knauff, Markus, editor. | Spohn, Wolfgang, editor.
Title: The handbook of rationality / edited by Markus Knauff and Wolfgang Spohn.
Description: Cambridge : The MIT Press, 2021. | Includes bibliographical references and index.
Identifiers: LCCN 2020048455 | ISBN 9780262045070 (hardcover)
Subjects: LCSH: Reasoning (Psychology) | Reason. | Cognitive psychology. | Logic. | Philosophy of mind.
Classification: LCC BF442 .H36 2021 | DDC 153.4/3—dc23
LC record available at https://lccn.loc.gov/2020048455