This is a section of doi:10.7551/mitpress/14922.001.0001

Open Minded

Searching for Truth about the Unconscious Mind

By: Ben R. Newell, David R. Shanks

Citation:

Open Minded: Searching for Truth about the Unconscious Mind
By: Ben R. Newell, David R. Shanks
DOI: 10.7551/mitpress/14922.001.0001
ISBN (electronic): 9780262375375
Publisher: The MIT Press
Published: 2023

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

7 One, Two, or More? System(s) of Thinking

> People make most judgments and most choices automatically, not deliberatively: we call this "thinking automatically."
>
> Over the past few decades, evidence has mounted that automatic thinking cuts across wide swathes of human behavior to the point that it can no longer be ignored. The anomalies . . . are not minor and scattered. They are systematic regularities that can be of first-order importance for health, child development, productivity, resource allocation, and the process of policy design itself.
>
> —*World Bank Report: Mind, Society and Behavior* (2015)

Consider the following question:

If a bat and a ball cost $1.10 in total and the bat costs $1 more than the ball, how much does the ball cost?

If you are like many people, your immediate answer would be "10 cents." You'd be wrong. Think a little more, and you'll see why.[1] This fluent but incorrect answer is supposedly the product of the kind of automatic thinking described in the opening quote from the World Bank's *Report into Mind, Society and Behavior*. The fact that the World Bank is concerned about this kind of thinking illustrates just how prominent and influential the notion of anomalous judgments arising from flawed, fast thinking has become. Similar statements can be found in reports from institutes of medicine, government departments, and countless consultancies spruiking their tools for overcoming the biases that arise from an overdependence on automatic thinking.[2]

At its core, dual-systems thinking appeals to the idea that mental processes can be dichotomized and compartmentalized into different boxes that house styles of thinking that in some way align or coalesce. Table 7.1 provides an illustrative example of the way in which mental processes get

Table 7.1
An example of dual systems of thinking.

System 1	System 2
Pragmatic	Logical
Fast	Slow
High capacity	Capacity limited
Nonconscious	Conscious
Automatic habits	Controlled
Associative	Rule based
Emotional (hot)	Unemotional (cold)
Independent of cognitive ability	Correlated with cognitive ability

Note: Attributes of systems are clustered according to the properties of apparently distinct and discrete modes of thinking. The evidence in this chapter suggests such dichotomies are simplistic at best and potentially harmful for understanding how the mind operates.

lumped together in these frameworks.[3] Broadly, system 1 captures the kind of automatic thinking that the World Bank is concerned about. System 2 is the slower, deliberative system, often caricatured as *homo economicus*—the rational, omniscient individual.

Debate about dual systems often appeals to intuitions and philosophical speculation, but these discussions typically avoid detailed reflection on concrete evidence. There is in fact a wealth of evidence that researchers have collected and interpreted over recent years as providing firm support for the existence of systems 1 and 2. We will certainly not attempt to review all this evidence here but will go into two examples in some detail. This is important: we should never build our theories about the mind purely on intuition, because on close examination, many intuitions can be revealed as mistaken. One of these examples concerns automaticity and the idea that some mental processes, such as accessing a word's meaning from its printed form, take place without the need for conscious voluntary control and cannot be prevented. But first we will look at evidence that logical thinking can be separated from intuitive thinking. The following example necessarily requires a rather extensive explanation, but the case allows some absolutely fundamental issues and assumptions about whether the mind is composed of two (or more) distinct systems to be investigated.

Belief Bias?

The case is based on a highly influential study conducted by a prominent advocate of the system 1/system 2 distinction and expert in human reasoning, Jonathan Evans, and involves a famous example of apparent irrationality in thinking, *belief bias*.[4] This refers to the common finding that when we judge the soundness of an argument, our judgments tend to be biased by how well the conclusion matches our prior beliefs and expectations. We all believe that ostriches can't fly so if given the argument:

All birds can fly.

Ostriches are birds.

Therefore ostriches can fly.

we judge the argument to be unsound. In fact the argument is perfectly valid; it's the premise that "all birds can fly" that's at fault. Purely from the perspective of judging whether the conclusion is logically valid given the premises, the believability of the conclusion and whether it fits with our knowledge of the world is irrelevant. And clearly people can override this bias. We can all see, given suitable time for reflection, that the argument is perfectly sound. Belief bias is closely related to *confirmation bias*, the tendency to actively search out for information that matches our prior beliefs and expectations.

But what happens when we are unable to reflect adequately on the validity of an argument because our conscious, reasoning mind is unavailable or occupied with something else? This is the question Evans asked. Participants were required to judge the validity of syllogisms such as this, rather harder than the one above:

No healthy people are unhappy.

Some astronauts are unhappy.

Therefore some astronauts are not healthy people.

It was made clear to participants that their task was to judge whether the final conclusion followed logically from the two preceding premises. It is a task about logical validity, not about whether the conclusion happens to be true or false in the real world. The study of syllogisms like this goes back to Aristotle, and they can (as this one illustrates) be very difficult. In fact, it is valid. According to the first premise, the set of healthy people includes none who are also members of the set of unhappy people. The second premise says that some astronauts are in the set of unhappy people, so these can't be at the same time in the set of healthy people. Hence, at least some astronauts are outside the set of healthy people. It is worth reiterating again that logical validity has nothing to do with the actual truth or falsity of any of the statements. Clearly there are not many unhealthy astronauts in the world.

Although actual truth or falsity is irrelevant to validity, it is a factor that Evans manipulated in the experiment, as can be seen from the examples in table 7.2. Two of the syllogisms have quite believable conclusions, while the other two have relatively unbelievable ones. At the same time, the logical validity of the conclusions was varied, two being valid and two being invalid. For example, the syllogism in the bottom left of the table (*no healthy people are unhappy*; *some astronauts are unhappy*; *therefore some healthy people are not astronauts*) is invalid because there is nothing in the premises that rules out the possibility that all healthy people are astronauts, in which case the conclusion does not hold.

The other crucial factor that Evans varied was time pressure: for one group of participants, there was no time pressure to make their decisions, while a second group had to make their validity judgments within five seconds. The key findings are shown in figure 7.1. For participants who completed the task under no time pressure, both validity and believability affected their likelihood of judging a syllogism as valid. This is shown in the solid black bars, which indicate that the validity of the arguments made a substantial difference to the likelihood of their being endorsed (the second and fourth bars from the left—the valid ones—are higher than the first

Table 7.2
Examples of the syllogisms used by Evans and Curtis-Holmes (2005)

	Believability of conclusion	
Validity	Believable	Unbelievable
Valid	No astronauts are unhappy. Some healthy people are unhappy. Therefore some healthy people are not astronauts.	No healthy people are unhappy. Some astronauts are unhappy. Therefore some astronauts are not healthy people.
Invalid	No healthy people are unhappy. Some astronauts are unhappy. Therefore some healthy people are not astronauts.	No astronauts are unhappy. Some healthy people are unhappy. Therefore some astronauts are not healthy people.

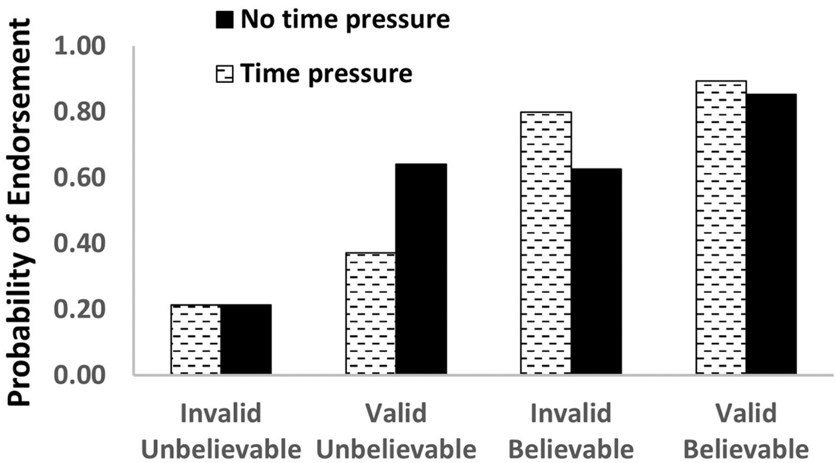

Figure 7.1
Results from the experiment by Evans and Curtis-Holmes. Participants read the syllogisms in table 7.2 and for each one decided whether the conclusion necessarily followed from the premises. In this graph, the vertical axis shows the average rate at which participants endorsed the conclusion of each type of syllogism as valid. The syllogisms were either valid or invalid, and their conclusions either believable or unbelievable. When decisions were made under no time pressure (solid bars), both validity and believability influenced choices. Under time pressure (hatched bars), in contrast, validity had only a very small effect on decisions and believability a much larger effect. At face value, this pattern seems to suggest that there are two different mental processes or systems that contribute to logical reasoning. The slow, logical system 2 dominates when ample time is available, but this system makes a smaller contribution under time pressure. The fast system 1, which promotes the intuition that believable conclusions are likely to be valid, dominates under time pressure.

and third—the invalid ones). In addition, believable arguments (third and fourth bars from left) were judged stronger on average than less believable ones (first and second bars).

Strikingly, time pressure had opposite effects on these influences. Now participants' judgments were less affected by validity but more affected by believability. This can be seen in the hatched bars in the figure. Validity had only a very small effect on the endorsement rate (the heights of the first and second hatched bars are similar, as are the third and fourth despite differing in validity), while believable syllogisms (whether valid or invalid) were much more likely to be judged acceptable than unbelievable ones (the third and fourth hatched bars—believable—are much higher than the first and second—unbelievable).[5]

The line from this pattern of results to the system 1/system 2 distinction is a straightforward one and is illustrated in the upper (dual-process account) section of figure 7.2. When there is no time pressure, the slow, logical system 2 is able to evaluate the strengths of the arguments using rational principles to distinguish valid from invalid deductions. System 1 also gets in on the act, creating a modest draw toward the more believable conclusions. Under time pressure, on the other hand, system 2 cannot operate as it normally would; hence, logical validity determines participants' judgments to a lesser extent, and the believability bias is enlarged as system 1 is "unleashed" to strongly influence responses.

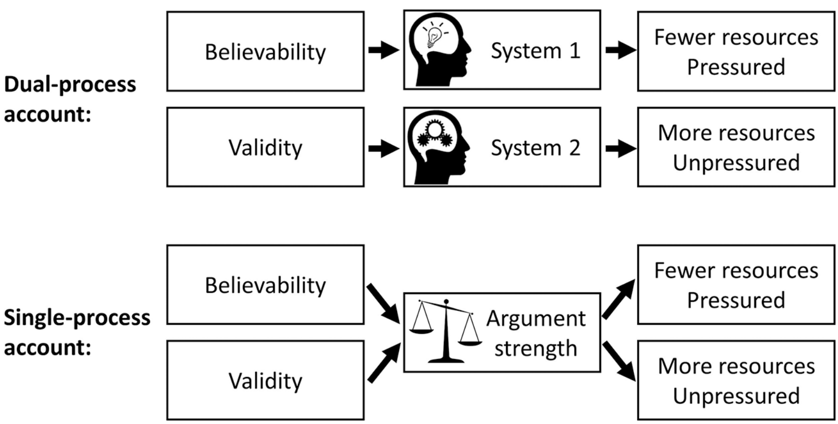

Figure 7.2
Two contrasting theories for explaining the findings of Evans and Curtis-Holmes. The dual-process account assumes the existence of a fast, automatic, and unconscious system 1 and a slow, conscious, and logical system 2. System 1, which dominates under time pressure, is intuitive and hence leads to choices that are heavily influenced by the believability of the conclusion of the syllogism. System 2 plays a large role when ample time is available, is logical, and leads to choices that are influenced by validity. The contrasting single-process account assumes that there is a single continuum of argument strength whereby valid syllogisms with believable conclusions are strongest and invalid ones with unbelievable conclusions are weakest. Both believability and validity hence contribute to this single measure of argument strength. Different mappings of argument strength onto behavior are responsible for the dominance of believability under time pressure and the greater influence of validity when time is available.

Observation and Inference

What is wrong with this account? A fundamental problem in psychology is the challenge of inferring the nature of the mind from outward behavior. Psychology is often defined as the science of behavior, a concept that emphasizes that observable behavior is to a large extent all we have to go on if we want to build theories about mental processes. In some aspects of perception and action, we might be able to measure activity in specific parts of the brain that are strongly associated with the content of our thoughts. For example, we can identify very specific brain regions associated with the perception of different colors or sounds and can use this knowledge to read the mind. From information about brain activity, it is possible to infer what underlying mental state the person is in, such as looking at a red or green patch. But these cases are the exception rather than the rule. They relate strongly to what we might think of as input or output processes in the brain. When it comes to thought more generally, we have to infer mental states. We cannot directly observe mental processes associated with whether some ongoing mental calculation is correct or how close a person is to reaching a difficult decision.

So the underlying states we refer to in explaining behavior, such as beliefs, desires, attitudes, and feelings, are hypothetical entities. Much of psychology and cognitive science is dedicated to the process and methods by which these "latent" states can be characterized from their effects on behavior, but there will inevitably be uncertainty about their relationship to behavior. Consider the example of two different ways of measuring knowledge, via speed and accuracy. One way of measuring individuals' general knowledge would be to give them unlimited time to answer a set of, say, fifty trivia questions. This method emphasizes accuracy of knowledge, with speed being irrelevant. Another way would be to give everyone a limited amount of time and measure how many questions they're able to answer correctly in that time. This method places more emphasis on speed of accessing one's knowledge. We might find somewhat different patterns from these measures. Person A might score much higher than B when unlimited time is available, but they might perform the same under time pressure. Person C might score better than D under time pressure but perform the same with unlimited time. It is, in short, difficult for us to know very much about the precise way in which underlying mental states map onto observable behavior.

For our psychological theories to be useful and falsifiable, there have to be some constraints on these mappings. A very strong constraint would be that they are linearly related, such that an X percent increase in the latent state (such as the amount of general knowledge) always maps onto a Y percent increase in our measurement of that knowledge. Under such a linearity assumption, this X-Y relationship would apply whether the person's general knowledge is poor or excellent. A much weaker constraint would be *monotonicity*, the milder assumption that if the latent state increases in strength, the observable behavior should either remain constant or increase but never decrease. This assumption captures the commonsense idea that as a person becomes more knowledgeable, she should always score at least as well, and never worse, on a measure of that knowledge.

The implausibility of linear mappings between mental states and behavior can be seen very easily by considering all-or-none actions. Think about your decision to take an umbrella with you when you leave home in the morning. You presumably have some strength of belief, between 0 percent and 100 percent, that it will rain today. Perhaps you've looked at the weather forecast or seen gray clouds out of the window. If we accept that your underlying belief lies on a continuous scale, then the linearity assumption says that your behavior—whether to take an umbrella—must also vary continuously in strength. But this is impossible: you either take an umbrella or you don't. In this case, it's obvious that the mapping is a step function. Between complete belief that it won't rain and some lower level of belief strength, say 30 percent, you choose not to take an umbrella, but as soon as your belief exceeds that level, you take an umbrella. Between that 30 percent level and certainty (100 percent), there is no behavioral indicator of your strength of belief. Your observable behavior (taking an umbrella) is identical whether the belief is 30 percent or 60 percent or 100 percent.

Latent States and the Implausibility of Dual Systems

How does this relate to the two-systems debate? The answer is that many of the most apparently convincing arguments for dual systems, including the dual-systems interpretation of Evans's findings, rely on extremely strong, and probably implausible, assumptions about mappings between latent mental states and behavior. If there is a linear mapping from underlying belief strength to endorsement rates, then Evans's results indeed suggest that two distinct systems or processes contribute to people's decisions. This can be seen most clearly in figure 7.1 by contrasting endorsement rates to the valid but unbelievable (VU) and invalid but believable (IB) syllogisms. Under time pressure, the IB syllogisms (third hatched bar from left) were more than twice as likely to be judged valid as the VU ones (second hatched bar), and hence to explain this difference we have to assume that the underlying belief strengths of these syllogisms differ greatly. However, if latent belief strengths map linearly onto behavior, this difference in the underlying strengths of the VU and IB syllogisms should also be evident when there is no time pressure, yet this is clearly not the case, as the figure shows (the second and third solid bars are almost identical in height). So it would be plausible to conclude that whatever is the latent process that is responsible for the difference in endorsement rates to these items, it is something that influences choices under time pressure but not when there is no time pressure. This is in effect what the dual-systems account proposes, with system 1 being the label for this process.

As soon as we relax the linear mapping assumption, this line of argument becomes much less compelling. Let us assume instead that the mapping from latent belief strength to endorsements has the form shown in figure 7.3, with valid and believable (VB) syllogisms having greatest strength, followed by IB, VU, and IU items in that order and with slightly different mappings when time pressure is or isn't applied. Both of these functions are monotonic in the sense that as strength increases across the four item types, endorsement either increases or at least remains steady. Crucially, an increase in strength is never accompanied by a decrease in endorsement rates.

With these mappings, Evans's data are easily explained. When there is no time pressure, validity is a major influence in that going from an invalid to a valid syllogism (from IU to VU, or from IB to VB) results in a large increase in predicted endorsements. At the same time, believability is also important: going from IU to IB or from VU to VB results in an increase in endorsements. So long as the syllogisms are ordered in the right way, the mapping is monotonic. In the contrasting case where decisions are made under time pressure, the mapping yields a different pattern in which believability is the dominant factor. Going from an unbelievable to a believable syllogism (from IU to IB or from VU to VB) results in a large increase in predicted

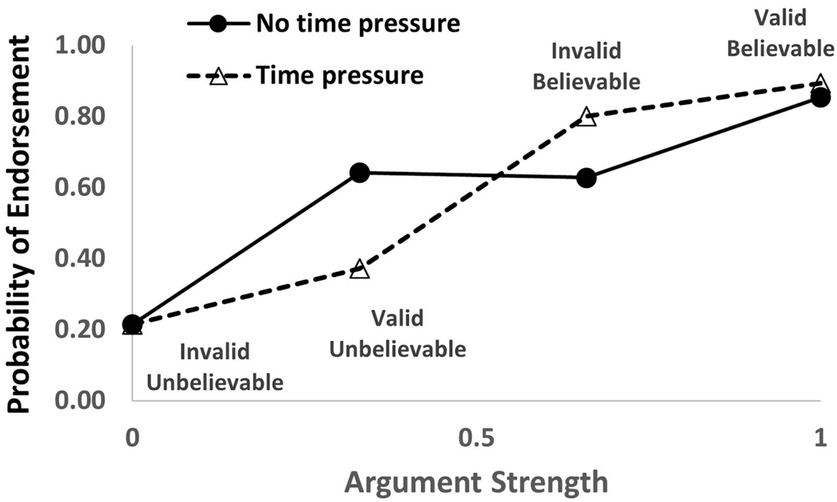

Figure 7.3
Hypothetical mappings from belief strength to behavior across the four types of syllogism that Evans and Curtis-Holmes used. The data are identical to those in figure 7.1, but the horizontal axis now orders the four syllogism types according to their hypothetical latent strength, with valid and believable (VB) syllogisms having greatest strength, followed by invalid believable (IB), valid unbelievable (VU), and invalid unbelievable (IU) ones in that order. The vertical axis indicates the probability that each type of syllogism is endorsed as valid. If slightly different functions are assumed for the time pressure and no time pressure conditions, Evans and Curtis-Holmes's results are perfectly reproduced, as indicated by the symbols on each line, which exactly mirror the results shown in the earlier figure. Crucially, both of the lines are monotonic in the sense that they always increase or stay flat as argument strength increases, and they never go down.

endorsements, whereas validity has a smaller impact, in that going from IU to VU or from IB to VB results in only a small increase in endorsements. Again, as long as the syllogisms are ordered in the right way, the mapping is monotonic.

In short, the results of this influential experiment only provide support for the dual-systems theory if a very strong assumption is made about the mapping from the underlying latent argument strengths to behavior—namely, that it is linear. But we have no evidence that it has this form. On the much weaker and more plausible assumption that this mapping is very unlikely to be perfectly linear, nothing in the results requires the postulation of two separate systems.

The astute reader may be wondering how this single-system account could ever be falsified. The answer is that any pattern of findings that cannot be reproduced by monotonic mappings would seriously question it. Suppose that we obtained data revealing that the different syllogisms cannot be ordered monotonically in the two conditions, time pressure and no time pressure, by any increasing measure of syllogism strength. For instance, imagine that Evans had found that endorsement rates to the VU syllogisms were greater than to the IB ones under conditions of no time pressure. This would mean that there is one ordering of endorsement rates (IB > VU) under time pressure but a different ordering (VU > IB) under no time pressure. There is no ordering of these items by latent strength that can generate such a contradictory pattern, and hence the single system account would be demonstrably inadequate, even with its very weak monotonicity assumption. However, extensive investigations of very many experiments have failed to throw up convincing evidence of such patterns.[6]

This reinterpretation of how people solve logical riddles casts serious doubt over the plausibility of dual-systems accounts. But are some simpler mental processes truly automatic in the sense of being completely outside our voluntary control?

Stroop: A Paradigmatic Automatic Effect?

In 1935 a young American scholar by the name of John Ridley Stroop published a paper entitled "Studies of Interference in Serial Verbal Reactions" that was to have a profound and enduring impact on psychology.[7] The phenomenon described in that paper now bears the author's name and is known to graduates of psychology programs the world over: the Stroop effect. The effect, which we will come to in a moment, has become synonymous with the idea of conflict between automatic and effortful, controlled processes. It is often taken as prima facie evidence of the mind's internal struggle between dual monolithic systems of cognition: one unconscious and outside voluntary control (system 1), the other conscious and controllable (system 2).

Perhaps surprisingly, neither the words *automatic* nor *controlled* appear anywhere in Stroop's original paper, and there is certainly no mention of dual systems, or consciousness. It is of course impossible to know what J. R. Stroop would have thought of his effect becoming so influential in the development of dual-systems thinking—he died in 1973, long before the current zeitgeist for dichotomies—but there are some clues. Thirty years after his original paper was published, he claimed to have no interest in the task he had created.[8] The "Stroop Scholar" Colin MacLeod recounts that Stroop's son told him that his father's psychological research was "insignificant to his Bible-oriented life and teaching."[9] Indeed, Stroop became a biblical scholar of some standing, preferring that calling to studying the apparent inherent conflict between mental processes.

The basic Stroop phenomenon is incredibly easy to demonstrate. Consider the examples presented in figure 7.4. For each panel, the goal is to read out loud the color of the ink that the word (panels a and b) or single letter (panels c and d) is written in. (For the purposes of this grayscale illustration here, you will have to imagine that the dark gray ink is blue and the light gray ink is yellow). Starting with panel a, you'll find it is pretty easy (and quick) to say "blue" and "yellow," but for panel b, it is trickier. There

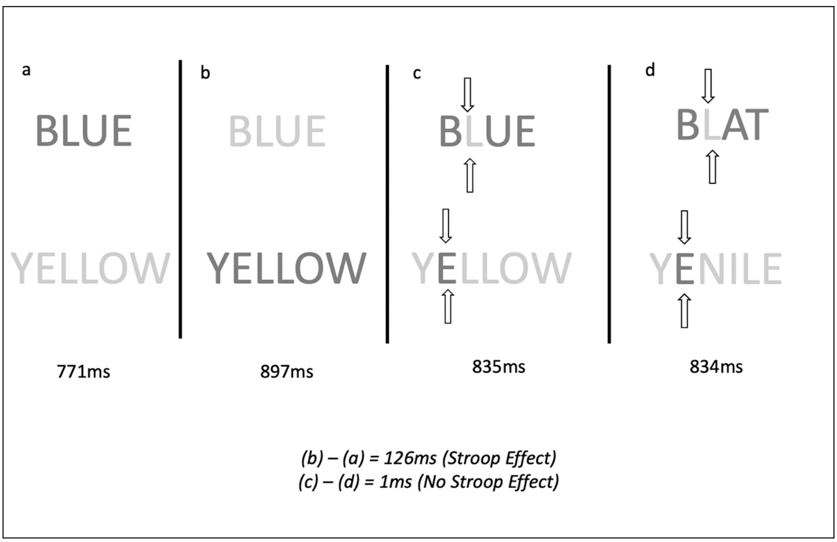

Figure 7.4
A demonstration of the Stroop effect and its disappearance. The task in each panel is to name the color of the ink that the word or cued letter is printed in. The difference in reaction times between panels a and b is often taken as evidence for the automaticity of word reading. However, finding that reaction times are identical in panels c and d casts doubt on this interpretation. (Note you need to imagine that the dark gray ink is printed in blue and the light gray in yellow.)

is a strong temptation to say "blue" instead of "yellow" for the first word and "yellow" instead of "blue" for the second. This difference in the ease of naming the color between panels a and b is the standard Stroop effect. In this example, on average, participants took 126 milliseconds longer to identify the correct color in the incongruent (panel b) compared to the congruent (panel a) case.[10] (An interesting, if admittedly arcane aside here, is that although the comparison of naming times in the congruent and incongruent versions is the standard Stroop effect, Stroop himself never actually gave participants the "congruent" version. Instead, he compared how long it took participants to read words versus name colors given lists of incongruent words like those in panel b.)

So why the difference in reaction times? The standard story is that the incongruent version takes longer because word recognition is automatic: lexical (determining the pronunciation) and semantic (determining the meaning) analyses of words are inevitably triggered by the presentation of a word. We cannot help but read the word and access its meaning, "BLUE," in the top line of panel b, and this triggers a voice inside our heads saying, "Say blue! Say blue!" Our controlled, conscious voluntary intention to say "yellow" is at the mercy of the all-powerful automatic process. But is this simple dichotomy, this conflict between systems, the whole story?

Consider now panels c and d of figure 7.4. Here, the task is to read out the color of the letter that is pointed to by the white arrows. Give it a try. In panel c, the color of the cued letter is incongruent with the whole word (the letter L is yellow, but the word says blue); in panel d, however, neutral "nonwords" have replaced the color words, but again a single letter is colored. Researchers often use these kinds of nonwords as controls to isolate particular aspects of visual and cognitive processing. The complete absence of a difference in reaction times between panels c and d is telling. If the mere presentation of a word inevitably and automatically triggered recognition and the processing of meaning, then we'd expect color naming to be slower in panel c than in panel d. Why? Because the word "blue" has a meaning that if processed automatically should interfere with our ability to say "yellow" in response to the cued letter L. In contrast, "BLAT" is meaningless and thus should not cause us any problems when naming the yellow L. The same goes for the blue E in "YELLOW" and "YENILE." The fact that we do not see any evidence for interference—the reaction times are almost identical—strongly suggests that word reading was not automatically triggered via the setup in panel c.

A better explanation for what happens in panels c and d is that we are able to exert some control over the deployment of our attention. Specifically, the arrows cue us to focus on a particular part of the word, thereby suppressing the word recognition processes that would otherwise lead to interference and slow naming speed. The words "BLUE" and "YELLOW" are right in front of our eyes to exactly the same extent in panels b and c, but we don't read them to the same extent: in the panel b setup, we do read them and suffer from Stroop interference, whereas in the panel c setup, we don't. So word reading cannot be an automatic process.

In fact, all kinds of variants of the Stroop task provide evidence that is more consistent with the idea that interference effects are graded rather than binary. Simply adjusting the ratio of congruent and incongruent words in a list affects the size of the Stroop effect. When congruent words dominate, the interference effects are larger because the overall context of the experiment encourages reading the word (rather than naming the ink). If word recognition were purely automatic, the ratio should not matter.[11]

Pragmatism versus Accuracy

Dual-systems thinking is deeply ingrained in contemporary discussion about the mind, as our earlier illustrations exemplify. How should we regard this framework in light of the kinds of counter-evidence we have just described? Daniel Kahneman stated in his influential book *Thinking, Fast and Slow*, "I must make it absolutely clear that [system 1 and system 2] are fictitious characters" that nonetheless, in his view, serve a useful purpose.[12] That pragmatic purpose is to make it easier to communicate complex ideas about the mind. The logic here is that it is comparatively easy for the general public to comprehend the notion that the mind can bring two quite distinct modes or systems of thought to bear in tackling a decision problem. Those two systems can be contrasted via a series of binary features such as intuitive versus rational, fast versus slow, automatic versus deliberate, hot (emotional) versus cold (unemotional), and so on (see table 7.1).

Such a framework can of course be helpful so long as it is accurate and there is reasonable consensus about the features of the two systems. But it will become a positive hindrance if the features are poorly defined and, worse still, if it turns out to be unsupported by the evidence. Scientists have sacrificed forests in proliferating variant dual-system models, each with its own characteristics. And in devising these models, which more and more resemble the epicycles devised by Greek astronomers to rescue the theory that the Earth revolves around the sun, their contact with hard facts becomes more and more distant. It is very rare to see proponents of the dual-systems approach provide explicit details of the predictions their models make that could be subjected to experimental tests.

Each of the individual dimensions mentioned above, such as automatic versus deliberate and emotional versus unemotional, is of great importance in understanding the mind, but the dual-systems approach provides very little explanation of why these dimensions should align. Why can we have automatic, emotional, or deliberate unemotional thinking but not deliberate emotional thinking? Our view is that behavioral scientists should first and foremost seek to understand the ways in which each of these dimensions is relevant to the mind. When, in the far distant future, we have a near-comprehensive grasp of them, then we can ask whether the dimensions tend to align or not. In our present state of knowledge, however, the degree of alignment assumed by dual-process accounts is little more than an untested assumption.[13]

There is also very little hard evidence that modes of thinking are binary rather than continuous. Both of the experimental effects we have discussed in detail in this chapter, logical reasoning and the Stroop effect, are better thought of from a continuous perspective. In the case of logical argument, we saw that the effects of validity and believability can be understood in terms of continuous argument strength. We also saw that whether a color word causes Stroop interference is a graded phenomenon. Many researchers now think of automaticity as a graded property, with different mental tasks placing demands on attention to varying degrees.[14]

To summarize, dual-systems thinking has become widely adopted not only among psychologists but more widely in debates about economic behavior, health, and public policy. This viewpoint may serve some useful communicative functions, such as conveying the important point that not all human decision making is based on logical or rational principles.[15] However, beyond this pragmatic function, the framework has a number of other implications, not all of them positive. It encourages binary thinking in places where it may not be appropriate, and it invites the view—for which there is very little evidence—that mental processes fall into clusters of aligned features.

This tension provides a segue into the second part of the book where we start to think in more depth about why the shaky theoretical foundations on which many claims about the unconscious mind are based have come to be ignored, forgotten about, or simply not known. (Oh, and in case you are still wondering, the answer to the bat and ball question is five cents.)

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: Newell, Benjamin R., 1972– author. | Shanks, David R.
Title: Open minded : searching for truth about the unconscious mind / Ben R. Newell and David R. Shanks.
Description: Cambridge, Massachusetts : The MIT Press, [2023] | Includes bibliographical references and index.
Identifiers: LCCN 2022038725 (print) | LCCN 2022038726 (ebook) | ISBN 9780262546195 (paperback) | ISBN 9780262375368 (epub) | ISBN 9780262375375 (pdf)
Subjects: LCSH: Subconsciousness. | Thought and thinking. | Self-consciousness (Awareness)
Classification: LCC BF315 .N479 2023 (print) | LCC BF315 (ebook) | DDC 154.2—dc23/eng/20230316
LC record available at https://lccn.loc.gov/2022038725
LC ebook record available at https://lccn.loc.gov/2022038726