3.2 Natural Logic

David P. O'Brien

Summary

Declarative memory requires a predicate–argument structure to record propositions, and this is provided by a natural logic. A natural logic requires a representational format, some inference procedures, and a program for implementing the procedures and must attend to pragmatic influences. I show how inferences from a natural logic are integrated in reasoning with inferences from other practical processes, although a reasoner is not aware of the various origins of their inferences. I argue that other theories that eschew any role for a natural logic are not adequate without being integrated with a natural logic.

1. The Historical Shift in Attitudes about Rationality and Logic

The pre-Socratic Greeks introduced the idea that humans are rational, and after Aristotle, the basic Western tradition identified human rationality with logic—a perspective that dominated the Western perspective until the end of the 19th century, with only an occasional challenge by the likes of Bacon, Hume, or Nietzsche. Indeed, Boole (1854) treated logic as the study of human mental acts, so logic and human thinking were understood as more or less equivalent and part of the essence of human nature. The transition from the 19th to the 20th century, however, brought challenges to this view as psychology separated from philosophy after the founding of the new science late in the 19th century.

Although the most obviously popular claim that humans are not basically logical is found in Freud (e.g., 1914), work by logicians also encouraged the separation of logic from human thought. Philosophers like Frege and Russell, for example, argued for the separation of logic and epistemology from psychology (e.g., Gabbay & Woods, 2009; George, 1997; Passmore, 1994), and as professional logicians have focused on finding a foundation for mathematics in logic with their emphases on complex proofs of soundness and completeness, they were not producing something inviting to someone seeking a model for ordinary human reasoning. At the same time, the newly established field of psychology was eager to divorce itself from philosophy (Boring, 1929; Mandler, 2007), seeking psychological explanations that were not imported from philosophical logic. So psychology turned away from logical rationality to rewards, punishments, reinforcers, instincts, defense mechanisms, and the like, and philosophers turned away from psychology to construct mathematical systems of logic that provided little of obvious interest psychologically. Henle (1962) noted, however, that these movements of drift between logic and psychology were based not on evidence but on a shift in attitude.

A shift in intellectual attitude is not an argument, and the fact that the assumption of human logical rationality had a two-and-a-half-millennia history should encourage us to reconsider its abandonment, keeping in mind that many different approaches have been proposed by logicians, and the lack of fit of any of them to psychology is not an argument that human reasoning lacks a logic. The stoic logic of Chrysippus, for example, was quite different from Aristotle's logic, and one might find things in Chrysippus that are psychologically interesting in ways that Aristotle's logic is not. The logics one finds in a contemporary standard-logic textbook, and in the professional journals for logicians with their truth tables and elaborate proofs of consistency and completeness, were not constructed to be models for human thought. To discover what kind of logic should be natural for humans, we need to think about why humans need a logic, what that logic would need to include, and what it would be unlikely to include.

2. Why We Should Expect a Natural Logic and What It Should Include

I begin with the fact that humans have a declarative memory, and it is axiomatic that to record something in declarative memory, a format must exist with which to record it (Braine & O'Brien, 1998; Fodor, 1981; O'Brien & Li, 2013). This format would need to keep track of entities and the properties that attach to those entities, and people hold propositional attitudes about what is expressed in the propositions that contain the entities and their properties, so the format must keep track of what is true and what is false so that people can assert, deny, suppose, question, attack, defend, believe, and doubt their propositions. Such a format is tantamount to a predicate–argument structure, so the approach described here begins with the assumption that the mind should include a logical format with which to record propositions. This logical format should bear some interesting relationship to the kinds of standard logics one finds in logic textbooks, although by no means would we expect it to be identical to any textbook logic, and we have referred to our approach as a natural logic or a mental logic as it is intended to provide an account of reasoning rather than to construct a formal system that is sound and complete. The task is different from that of a logician. Indeed, it begins by acknowledging that completeness of human logical reasoning has not been a consideration in human bioevolutionary development, although one would want their thinking to be sound in at least some meaningful way. The first task for a natural-logic theory is to describe the format, and once the format has been established, we need to address how inferences are drawn in the format, what sort of logical semantics is required, and how the natural logic integrates with inferences that stem from pragmatic concerns.

Standard-logic textbooks provide a language that includes symbols for connectives or operators (e.g., $\wedge$, $\vee$, $\neg$, $\supset$) and quantifiers (e.g., $\forall$, $\exists$), and such symbols correspond loosely to ordinary words of daily language (e.g., *and*, *or*, *not*, *if*, *all*, *some*). They also provide a truth-functional semantics in the form of truth tables, with *If p then q* given the truth-functional assignment of being true unless *p* is true and *q* is false—a definition that makes true such propositions as *I did not have a gun, but if I did, then I'm guilty as charged*, which few people would want to utter.

Propositions of this latter sort exemplify the paradoxes of material implication, and they provide a serious problem for an attempt to argue that the truth table for *if* can provide a sensible model for how people reason with *if*-propositions. Our natural-logic approach thus looks elsewhere to find the meanings of logic operators. Compare the following two propositions:

All the children who went to the zoo were in the third grade.

Some of the children who went to the zoo were in the third grade.

Suppose we discover that Álvaro went to the zoo, and we ask whether we can infer that Álvaro was in the third grade. The proposition beginning with *all* provides a basis for the inference that he was in the third grade, but the proposition beginning with *some* does not. Indeed, someone who does not understand which sentence sanctions the inference and which does not could not be said to understand the meaning of the words *all* and *some*. This illustrates how the natural-logic approach understands the logical semantics for its vocabulary. The meanings of the terms are given by the inferences that they sanction. The tradition of understanding semantics in terms of the inferences they sanction follows Wittgenstein (1958), Gentzen (1935/1964), Harman (1982), and Block (1998), among others, and has become known as a "conceptual-role" or "inferential-role semantics." This approach has been presented as a general theory for semantics across the lexicon, but our natural-logic approach makes no claims beyond the vocabulary of the natural logic, and for this part of the lexicon, a semantics based on the inferences that a term sanctions is straightforward because the schemas and reasoning program of natural-logic theory provide a simple and succinct description of what the forms are for those inferences.

I turn now to the question of the inference procedures, both because these procedures provide the logical semantics and, moreover, because an intelligent species needs inferential procedures that go beyond the information given (e.g., Bruner, 1997). Our hunter–gatherer ancestors would have benefited from an ability to make simple and immediate inferences. Imagine a gatherer in the forest who knows that the best locations for finding food at this time of the year are either next to the river or in the forest near the canyon. She encounters her cousin who has just returned from the riverbank and tells her that there is little to collect there this year as locusts have been there, so a simple and direct inference can be made that she is better served going to the forest near the canyon. Surely our ancestors needed to make such inferences to have survived and produced progeny, for it was their logical reasoning that kept them from needlessly visiting multiple sites in useless searches for food. This particular kind of inference can be captured by an inference schema for disjunctions, that is, alternatives, and the stoic logician Chrysippus attributed such a schema to dogs following a trail of scent as well as to humans. (For discussions of why immediate inference procedures are not provided by the way professional logicians deal with truth conditions, see Braine, 1978; Dopp, 1962; Fitch, 1973; Gentzen, 1935/1964; Lakoff, 1970.)

Simply having some inference-making schemas would not be sufficient: a reasoner would need some sort of reasoning program that applies the schemas to produce lines of reasoning. Further, given that our thoughts usually refer to practical concerns, an adequate theory of logical reasoning also would require an account of how the logical inferences are coordinated with inferences that refer to practical knowledge and practical interests. To exclude inferences that go beyond what is available on logic alone would be irrational on the part of investigators into rationality. An investigator of human reasoning should not interpret inferences made in laboratory reasoning tasks that do not follow from logic alone as irrational. Such inferences may well be quite rational even when not inferable on logical grounds alone.

This chapter presents the natural-logic theory of Braine and O'Brien (e.g., 1991, 1998) as a case study in how a natural logic can be developed. The presentation of the theory has three parts: the representational format (section 3), a set of reasoning schemas and a reasoning program that implements the schemas in constructing lines of reasoning (section 4), and a set of pragmatic principles that describe the interactions of the logical schemas with epistemic knowledge and practical goals (section 5).

3. A Representational Format for a Natural Logic

Standard-logic textbooks first introduce the sentential level, which includes atomic propositions and connectives for negation, conjunction, disjunction, conditionals, and so on. A predicate level adds quantifiers and a predicate–argument structure, which can make it appear that reasoning at the predicate level is more difficult, but natural logic need have no such expectation.

Constructing a natural logic at the sentential level requires fewer decisions than does the predicate-logic level, providing only connectives but not quantifiers and the structure that quantifiers entail. The connectives in a natural logic are not identical to those in a standard logic. For example, "If p then q" is not equivalent to the material conditional of $p \supset q$, which as discussed earlier makes "If p then q" true whenever p is false. So the natural logic presented in Braine and O'Brien (1998) eschews the symbols one finds in logic textbooks and represents sentential connectives with italicized English words: *if* for conditionals, *and* for conjunction, *or* for disjunction, and *not* for negation, although we understand that these words have additional pragmatic interpretations that go beyond their purely logical meanings, and in ordinary speech, words can convey meanings that differ from their ordinary surface meanings. A mother who tells a child, "Put your hand in the cookie jar and I'll send you to your room," is using *and* where one formally would expect the word *if*, and we would not think the proposition requires that "The hand is put in the cookie jar" and "The child will be sent to their room" are both true in order for the assertion to be true. Even though on the surface, the assertion uses a word usually reserved for conjunction, the meaning intended clearly is conditional. Meaning thus is not found in the surface structures of sentences but in the propositions that are held in a deeper language of thought. Linguistic input thus is translated into a language of thought from the language of speech, and then inferences made in the language of thought are translated back into the language of ordinary speech (see Braine & O'Brien, 1998; O'Brien & Li, 2013).

Constructing the representational format at the predicate-logic level requires that some decisions be made about the internal structure of propositions. Consider a universally quantified disjunction in a standard-logic textbook:

$(\forall x)(Px \vee Qx)$, "For all x, x satisfies P or x satisfies Q,"

where the universal quantifier is placed outside the proposition that is being quantified, and the quantifier has scope over all instances of the bound variables within the parentheses, whatever the content predicates are to which the variables are attached.

A corresponding natural logic representation is:

S_1[All X] OR S_2[PRO-All X],
"All the X's satisfy S_1 or they satisfy S_2,"

where the scope of quantification is inside the expression and is specific to particular content domains. Quantificational scope in this example is marked by a PRO-notation, referring to the pronoun in the second clause that is an anaphoric reflex to the quantifier in the first clause. This choice about how to represent quantification reflects the use of pronouns that is typical in natural languages as a way to convey quantificational scope. Ioup (1975) surveyed 14 languages and found no evidence for the type of outside scope that is typical in standard-logic textbooks;

instead, quantifiers tend to be lexicalized inside noun phrases, as they are in the natural-logic representations, and quantification is bound to the specific content, as it tends to be in natural languages.

One needs to distinguish between alternative interpretations of the surface grammar of sentences like "All the boys are carrying a box," where one interpretation is that there is a single large box that is being carried by a group of boys, and a different interpretation is that each boy is carrying a different box from the other boys. The first interpretation can be captured by a representation of this sort:

CARRY BOX (BOYS),

and the second interpretation can be captured by one of this sort:

CARRY BOX (All BOYS).

Such sorts of collective sets are not at all unusual. Braine and O'Brien (1998) provided the example

> "The prosecutors convinced the 12 members of the jury of the defendant's guilt,"

CONVINCED (PROSECUTORS) (All JURORS),

where the collective group of prosecutors convinced each individual jury member.

4. Inference Schemas and a Reasoning Program

Natural-logic theories propose that people construct lines of reasoning governed by inference schemas (e.g., Braine & O'Brien, 1991, 1998; Fitch, 1973; Gentzen, 1935/1964; Rips, 1994), and Braine and O'Brien emphasized the need to seek empirical evidence for guidance about which procedures used to construct inferences in lines of reasoning should be included. For example, Braine, Reiser, and Rumain (1984) found that college students tend not to accept "*p* or *q*" as following from *p* alone, even though it is a valid inference in standard logic and is included in the natural logic of Rips (1994); it was excluded on empirical grounds by Braine and O'Brien (1998).

Braine and O'Brien (1998, tables 6.1 and 6.2) presented a set of 14 schemas at the sentential level together with a reasoning program that applies the schemas to construct lines of reasoning. I describe these in some detail here to provide a sense of what is involved. Seven schemas are called "core schemas," and they are applied without restriction by the reasoning program whenever their conditions for application are met. For example (presented here in simplified form), when one holds "*p* or *q*" and not-*p* together in working memory, one infers *q*, as did our gatherer described earlier who decided to go to the area of the canyon to procure food. Another core schema is modus ponens, which holds that when one considers "If *p* then *q*" and *p* jointly, one infers *q*.

Two schemas are called "feeder schemas," and their application is restricted so that they are applied only when their output can feed another inference. One of these infers "*p* and *q*" when both *p* and *q* are in working memory, and the other infers *p* alone (or infers *q* alone) from "*p* and *q*." The restriction is required because if left unrestricted, these feeder schemas could produce infinite loops. Also, people do not report their output when asked to write down everything that follows from some premises, yet such schemas seem to exist because people apply, for example, modus ponens to "If *p* and *q* then *r*" when they have both *p* and *q* separately.

The final four schemas include two that pertain to incompatibility and two that pertain to supposition. Supposition occurs when a proposition is treated as true for the purposes of reasoning with the supposed proposition, even when one does not know whether or not it is true. One of the supposition schemas combines with an incompatibility schema so that the supposition can be falsified as in a *reductio ad absurdum* argument. The other is a schema for conditional proof, which is used to infer propositions of the form "If *p* then *q*" or to evaluate whether "If *p* then *q*" follows from a set of premises. To infer a proposition of the form "If *p* then *q*," first suppose *p*. When a set of premises and background assumptions together with the supposition of *p* lead to the inference of *q*, this schema provides the inference of "If *p* then *q*."

The schemas are applied by a reasoning routine, which has two parts: a direct-reasoning routine (the DRR) that we claim is universally available and applied automatically, and some reasoning strategies that go beyond the DRR and are widely available among college students. The schema for conditional proof has a procedure in the DRR that applies when a line of reasoning is constructed to evaluate "If *p* then *q*": the proposition *p* is added to the premises as a suppositional premise, and reasoning proceeds toward *q*. When *q* is derived, "If *p* then *q*" is evaluated as true. When not-*q* is derived, "If *p* then *q*" is evaluated as false. The latter evaluation would not be made in standard logic with its material conditional—because *p* might be false, even when supposed as true, although both young children and adults seem to use this procedure (Braine & O'Brien, 1998, chapters 7 and 17).

Note that the combination of the schemas and the reasoning program enables precise predictions about

This is a portion of the eBook at doi:10.7551/mitpress/11252.001.0001

what inferences people make, and some of these inferences differ from what would be predicted if natural reasoning were following standard textbook logic, like the evaluations of when conditional propositions are false just described. The combination of schemas and reasoning program also helps us understand the consistency of the schema for conditional proof and the modus ponens schema, which together provide the inferential semantics for *if*. Given that "If *p* then *q*" is derived by supposing *p* and then showing that this entails *q*, the introduction of *p* as a premise when one knows "If *p* then *q*" allows the conditional nature of *q* to be discharged so that *q* alone can be asserted. That is, modus ponens releases *q* from requiring the marking of *p* as suppositional when one knows that *p* is true, because the schema for conditional proof ensures that *q* is true when *p* is.

Reasoning under a supposition requires a constraint (see Braine & O'Brien, 1991). A sound line of reasoning can proceed only from true premises, and when a line of reasoning is done under a supposition, one needs to be sure that the propositions introduced under the supposition would still be true given the supposition. When it is not clear whether a proposition should be excluded under a counterfactual supposition, logic per se provides no clear way to resolve such cases, and interlocutors would need to negotiate about whether exclusion is required. Note, by the way, that the natural logic is able to deal with counterfactual conditionals without recourse to some special semantics that requires reference to a possible world that is most similar to the actual world. After all, what would the possible world be that is most close to the actual world when we consider the counterfactual "Philadelphia is north of New York City"? One would have to decide to move New York southward or to move Philadelphia northward, with all of the messy repercussions that such a decision would have.

For a detailed presentation of the complete set of predicate-level schemas and the reasoning program for their implementation, see table 11.3 in Braine and O'Brien (1998). I present the following schema for a universally quantified disjunction to illustrate the nature of the predicate-level schemas and to make the point that even though the notation is more complex than at the sentential level, the inferences are not necessarily more difficult.

$$S_1[\text{All } X] \text{ OR } S_2[\text{PRO-All } X];\ [\alpha] \subseteq [X];\ \text{NEG } S_2[\alpha] \therefore S_1[\alpha].$$

For example, "All of the boys from the village are at the river fishing or in the forest hunting. Uirá and Werá are boys from the village and they are not at the river fishing, so they are in the forest hunting." Unlike standard logic, which suggests that reasoning at the predicate level is intrinsically more difficult than at the sentential level, this example illustrates that intuitively, this is not the case in natural logic.

5. Pragmatic Principles

People do not reason only by applying natural-logic schemas. An intelligent person uses any information that is pertinent, and inferences thus can be made on the basis of many kinds of processes. A person should interpret words in plausible ways and not limit their interpretations to what a professional logician has in mind. *If*, for example, can invite additional sorts of interpretations (Geis & Zwicky, 1971)—for example, "If *p* then *q*" can invite interpretation as a biconditional, depending on context; if I am told that *if I paint the shed I will receive $50*, I expect that if I do not paint the shed, I will not receive the money. Natural logic does not constrain *or* as either exclusive or inclusive, but certain kinds of situations surely invite one interpretation over the other, which can expand the available inferences.

Story grammars or scripts provide additional inferences that a rational person would make. My knowledge about eating breakfast in a typical American diner tells me that I will not drink both tea and coffee, and this knowledge provided by a restaurant script can feed a natural-logic inference. Although a cognitive scientist is interested in the sources from which inferences are made, people are unlikely to attend to the varied sources that can lead to an inference or to distinguish purely logical inference from other sources. These things are not evidence of irrationality; in ordinary reasoning, inferences from various sources intertwine with inferences from natural logic to provide reasoning that meets practical concerns as lines of reasoning are constructed.

6. Empirical Support for the Braine–O'Brien Natural Logic

The most basic prediction of natural logic is that the inferences available on the DRR alone will be made routinely unless something complicates matters, and the evidence, generally from problems presenting neutral content, supports this prediction. Further, inferences that are valid in standard logic but not available on the schemas of natural logic are made far less often, as are inferences requiring strategies that go beyond what is available on the DRR alone. Finally, problem length per se does not predict problem solution, which depends for the most part on whether the required inferences are available in the DRR (Braine et al., 1984; Braine & O'Brien, 1998; see review in O'Brien, 2004).

Braine et al. (1984), with sentential problems, and Yang, Braine, and O'Brien (1998), with predicate-level problems, presented DRR problems and asked people to rate on a Likert-type scale the relative difficulty of each problem. Regression analyses provided weights for each schema. Problem difficulty was assumed to equal the sum of the difficulty weights for all of the schemas required to solve a problem. These weights were used to compute predicted difficulties for different problems that were presented to different people. The percentage of variance accounted for in correlations between the predicted and observed difficulties was 66% for sentential problems (55% with problem length partialed out) and 69% for predicate-level problems (56% with problem length partialed out). This indicates that people were not only solving these problems but also solving them using the natural-logic schemas.

When the schemas were embedded within story scenarios and readers were asked whether the natural-logic schemas were inferences or paraphrases, people showed that they did not know that they were making inferences at all, although inferences based on story scripts, grammars, and so forth were judged as being inferences rather than paraphrases (Lea, 1995; Lea, O'Brien, Fisch, Noveck, & Braine, 1990; O'Brien, Roazzi, Dias, & Soskova, 2007). Thus, the natural-logic inferences were made so easily that people did not think of them as requiring reasoning.

Braine et al. (1995) and O'Brien, Braine, and Yang (1994) presented problems with multiple premises and asked participants to write down all of their inferences in the order they made them. Consider the following problem with five premises from O'Brien et al. (1994) that referred to some letters written on a blackboard: (a) "N or P," (b) "Not N," (c) "If P then H," (d) "If H then Z," and (e) "Not both Z and S." Premises (a) and (b) trigger a logic schema to infer P, which then combines with (c) to infer H, which then combines with (d) to infer Z, which then combines with (e) to infer not-S. So, as each premise was read in turn, a schema was triggered that provided an inference. A matched problem presented the same premises but with the order reversed, so that no schema was triggered until all premises had been read and the information in the last two premises could trigger a schema. For both problems, the same inferences were made in the same order, demonstrating that the order of inferences was determined by availability of schemas. In summary, the predictions about which problems will be solved, about the relative difficulties with which problems are perceived, and about the orders in which inferences are made are supported by the available data. Further, when the predictions differ from what would be predicted by a standard textbook logic, the data support the predictions of natural logic. Supporting evidence that the core inferences are available in children is presented in Braine and Rumain (1983) and Braine and O'Brien (1998).

7. Why Theories of Reasoning Need a Natural Logic

Proponents of several other theories have argued against any role for a natural logic. O'Brien and Manfrinati (2010) and O'Brien and Li (2013) have provided detailed discussions to counter their arguments against a natural logic. I concentrate here on whether those theories are adequate without including a natural logic.

Cheng and Holyoak (1985) introduced some content-specific inference rules for permissions and obligations that they proposed are inductively acquired, and Cosmides (1989) introduced some social-contract rules that she proposed have been acquired through our bioevolutionary history. Both Cheng and Holyoak (1985) and Cosmides (1989) argue that people do not use a natural logic but only content-dependent processes. Both theories rely on variants of Wason's selection task, comparing performance with universally quantified indicative conditionals to performance with sentences that should trigger either social-contract rules or pragmatic schemas. The problems required people to identify situations that could falsify the indicative conditionals or find rule violators of the pragmatic or social-contract rules, that is, instances of "*p* and not-*q*." Both theories reported more correct answers with their content, which they argued indicates that people do not reason with a natural logic but with their content-specific processes.

I address here whether these two theories provide an adequate account of human reasoning without including a natural logic, given that they specifically argue against any natural logic. (O'Brien & Manfrinati [2010] and O'Brien, Roazzi, Athias, & Brandão [2007] provide more extensive discussions of difficulties in interpreting the data for these content-specific theories.) The contention of both theories that reasoning relies not on a content-general natural logic but on some content-specific processes is based on very little evidence. They have investigated only a handful of conditionals that express permissions, obligations, or social contracts, and to support the notion that content-specific processes are sufficient to replace natural logic, they would need to provide a much larger set of types of content. I know

This is a portion of the eBook at doi:10.7551/mitpress/11252.001.0001

of no metric with which to estimate the number of content areas that one would need to provide a general reasoning theory, but it must be immense, and I find it hard to imagine that either inductive or bioevolutionary methods would have provided such a number of content-dependent modules. Further, they would need to present evidence beyond a sole reliance on Wason's selection task.

O'Brien, Roazzi, Dias, Cantor, and Brooks (2004) and O'Brien (1995) discussed why data from Wason's selection task are difficult to interpret, and there is little discussion by critics of natural logic of exactly how one would solve the task using logic. The selection task actually is a meta-logical task when presented with indicative conditionals—it requires judging the conditions under which the truth of a conditional could be tested, so a logic inference alone would not be sufficient, whereas the task presented with pragmatic or social-contract conditionals and requiring only identification of rule violators is not a meta-logical task. No attention has been given to this difference by proponents of these theories. If these content-specific theories are to expand their focus, they will need to present problems other than the selection task.

Finally, these theories need to explain why *if*-sentences are used to convey these regulations. Let me illustrate the problem by pointing out an implicit difference in developmental expectations between the pragmatic-schemas approach and the natural-logic approach. If one assumes that the pragmatic rules are inductively acquired, as pragmatic-schemas theory does, then a content-general schema of the sort we describe would enter late in the process only, after many pragmatic schemas had been acquired, yet Braine and O'Brien (1998) present evidence from spontaneous speech that young children grasp the schemas for conditionals quite early.

Oaksford and Chater (e.g., 2010) proposed that the probability that someone will accept an inference from a conditional premise depends on their conditional belief in the conclusion given the premises and other background information. They make several arguments against natural logic, but their arguments, especially their arguments about defeasibility, stem from treating natural logic as equivalent to standard logic. More important is the question of whether Oaksford and Chater are able to account for the data without including a natural logic. Their approach depends on prior beliefs about the premises on which inferences rely, so the approach is mute when people receive premises about which they have no prior beliefs. The approach thus cannot address the large body of evidence in favor on the natural-logic approach that refers to materials about which participants have no prior beliefs, like the size and colors of beads in an urn or letters written on an imaginary blackboard. Further, Oaksford and Chater have addressed only Wason's selection task and the four conditional syllogisms, and they need to explain the larger set of problems that we have investigated, which include conjunctions, negations, disjunctions, and conditional conclusions.

Evans and Over (2004) and Over and Evans (2003) proposed that people evaluate conditionals by imagining two situations, one in which both p and q are true and one in which p is true and q is false, and then assessing the probabilities of the two cases, and the probability of "If p then q" is assigned from the comparison of probabilities for the two models. So, "If p then q" is functionally equivalent to $P(q \mid p)$ computed from the probabilities assigned to the two models.

This proposal bears a clear resemblance to the natural-logic schema for conditional proof in that it begins with the supposition of p in constructing the two models in which p is true and then reasons toward q or not-q, but Evans and Over (2004) argued that because the natural-logic approach does not include either a possible-worlds semantics or a mental-models semantics, it cannot account for these probability judgments. Why one would need a possible-worlds or a mental-models semantics to judge these probabilities is not clear.

To illustrate that one does not need a separate logical semantics to make such judgments, imagine two boxes, a red box with 80 beads and a blue box with 20 beads. You are told that I have a bead in my hand and are asked to judge the probability that *if the bead came from one of these boxes, it came from the red box*. Using the natural-logic procedures for the conditional-proof schema, you begin by supposing that the bead came from one of these boxes, and you reason toward the proposition that it came from the red box. You discover that you cannot derive whether or not it came from the red box, so you know that the conditional cannot be shown to be true or false, but given that there are four times more balls in the red box and you are asked to judge the probability of the conditional, you interpret it as asking what is the probability that the ball came from the red box and respond that the probability is .80. No independent logical semantics of any other kind was required for someone to give such a response.

The theory of Evans and Over is not adequate to account for reasoning in general without a natural logic because their theory has not been developed beyond a single kind of judgment about conditionals and thus

This is a portion of the eBook at doi:10.7551/mitpress/11252.001.0001

cannot provide an account of the larger set of findings concerning other connectives. It has the advantage over the content-dependent theories because it is not tied to a single content domain, and it obviously bears a close resemblance to the Braine–O'Brien natural-logic theory. Integrating their proposal with a natural logic thus seems eminently possible, and such an integration clearly needs to happen if their theory is to address a more general class of reasoning types, although the natural-logic theory seems already capable of making the judgments their theory describes.

I turn finally to mental-models theory (e.g., Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991, 2002)—the only other theory to address the same broad class of reasoning problems that natural logic addresses—whose proponents consistently deny that there is any role in human reasoning for any natural logic. The models theory proposes that humans construct models that are adaptations of truth tables but represent not truth assignments but possible states of affairs. Because working memory is limited, people usually construct only some of the possible states of affairs. Although a complete representation of "If *p* then *q*" includes the three models corresponding to the material conditional, it usually is represented with

[*p*]   *q*
. . .

where the ellipsis acts as a reminder that additional models could be constructed, and the brackets around *p*—called "exhaustivity markers"—constrain any such additional models so as not to include a token for *p* without a token for *q*.

If the minor premise for a modus ponens argument is added to the model for "If *p* then *q*," one gets

[*p*] *q*;   *p*;   *p* *q*
. . .

This set of models separates models for the two premises and then their combination with semicolons, and the conclusion, *q*, is "read off" from the final model and does not include *p* presumably because pragmatics precludes inclusion of categorical premises in a conclusion. In this way, they assert, one can derive a modus ponens conclusion with only the manipulation of models and no logic.

Precluding any logic from their theory, including the variables in a predicate–argument structure, is a significant loss. Johnson-Laird and Byrne (1993) wrote that although models do not contain variables, variables do occur "in the initial semantic representations" (p. 376), that is, the representations from which they say their models are constructed.

Models for quantified propositions are almost identical to those for the sentential connective, and the complete model for "All *P* are *Q*" is

[*P*]   [*Q*]
[*P*]   [*Q*]
[~*P*]   [*Q*]
[~*P*]   [~*Q*]

Let's consider how models come from the premodels representations. Quantifiers provide "the raw material for a recursive loop that is used in building or manipulating a model. Thus, the universal quantifier 'all' elicits a recursion that deals with a set exhaustively, whereas the existential quantifier 'some' elicits a recursion that does not" (Johnson-Laird & Byrne, 1991, p. 178), which is illustrated with the example "All *x*'s are equal to the sum of some *y* and some *z*," which is parsed to yield

(All x) (Some y) (Some z) ($x = y + z$).

A model, then, is constructed with an arbitrary value for the first variable term (x), another arbitrary value for the second (y), and a constrained value (the degrees of freedom being exhausted) for the third variable (z), looping over the equation several times, recording the output each time, resulting in a model like

[8 6] (1 6 4 2) (7 7 2 2 4 4),

where the two numbers in the first set are equal to some number in the second set plus some number in the third set. In this, Johnson-Laird and Byrne have succeeded in their goal of constructing a representation that has no logical variables or quantifiers, but at what cost?

They have told us repeatedly that conclusions are "read off" from final models, so what conclusion could be read off here? Consider the simplest logic inference—reiteration: $p \therefore p$. But here, the meaning of the initial representation prior to construction of the model is gone, and one has no way of retrieving "All *x*'s are equal to the sum of some *y* and some *z*" from the final model. This is a steep cost to fulfill the goal of creating a reasoning theory that requires no logic inferences and no variables and only the manipulation of models. It's difficult to see that the procedure is analogous to the standard interpretation of quantifiers in the predicate calculus. The quantifiers and variables of the initial proposition have disappeared, and without them, one cannot retrieve the assertion from the model. Clearly, mathematical reasoning cannot be accomplished purely through the manipulation of models. How could one use this sort of procedure to prove "All natural numbers that end in zero are divisible by five" when one is limited to inclusion only of individual exemplars, square brackets to mark exhaustivity, and some ellipses to remind one that additional models might need to be constructed? Yet my students in undergraduate classes over many years have succeeded in understanding the necessary truth of this statement.

8. Conclusion

I have argued that an adequate account of human reasoning must include some predicate–argument structure and some inference processes for going beyond the information given and that other ways of thinking need to be integrated with the natural logic. Searching for counterexamples as one does with mental models would be a valuable skill to add to what is provided by natural logic, as would be the specialized searches for violators of pragmatic and social-contract rules. But to focus only on such specialized skills at the expense of the more general architecture provided by a natural logic would be to throw out the baby with the bathwater.

References

Block, N. (1998). Conceptual role semantics. In E. Craig (Ed.), *Routledge encyclopedia of philosophy* (Vol. 2, pp. 242–256). London, England: Routledge.

Boole, G. (1854). *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities*. London, England: Walton and Maberly.

Boring, E. G. (1929). *A history of experimental psychology*. New York, NY: Century.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, *85*, 1–21.

Braine, M. D. S., & O'Brien, D. P. (1991). A theory of *if*: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*, 182–203.

Braine, M. D. S., & O'Brien, D. P. (1998). *Mental logic*. Mahwah, NJ: Erlbaum.

Braine, M. D. S., O'Brien, D. P., Noveck, I. A., Samuels, M., Lea, R. B., Fisch, S. M., & Yang, Y. (1995). Predicting intermediate and multiple conclusions in propositional logic inference problems: Further evidence for a mental logic. *Journal of Experimental Psychology: General*, *124*, 263–292.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. Bower (Ed.), *Psychology of learning and motivation: Vol. 18. Advances in research and theory* (pp. 313–371). New York, NY: Academic Press.

Braine, M. D. S., & Rumain, B. (1983). Logical reasoning. In J. H. Flavell & E. Markman (Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 263–339). New York, NY: Wiley.

Bruner, J. (1997). *The culture of education*. Cambridge, MA: Harvard University Press.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

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.

Dopp, J. (1962). *Logiques construites par une méthode de déduction naturelle* [Logics built by a natural deduction method]. Louvain, Belgium: Nauwelaerts.

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Cambridge, England: Cambridge University Press.

Fitch, F. B. (1973). Natural deduction rules for English. *Philosophical Studies*, *24*, 89–104.

Fodor, J. A. (1981). *Representations: Philosophical essays on the foundations of cognitive science*. Cambridge, MA: Harvester Press.

Freud, S. (1914). *Psychopathology of everyday life*. London, England: T. Fisher Unwin.

Gabbay, D. M., & Woods, J. (Eds.). (2009). *Handbook of the history of logic: Vol. 5. Logic from Russell to Church*. Amsterdam, Netherlands: Elsevier.

Geis, M., & Zwicky, A. M. (1971). On invited inferences. *Linguistic Inquiry*, *2*, 561–566.

Gentzen, G. (1964). Investigations into logical deduction. *American Philosophical Quarterly*, *1*, 288–306. (Original work published 1935)

George, R. (1997). Psychologism in logic: Bacon to Bolzano. *Philosophy and Rhetoric*, *30*, 213–242.

Harman, G. (1982). Conceptual role semantics. *Notre Dame Journal of Formal Logic*, *23*(2), 242–256.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, *69*(4), 366–378.

Ioup, G. (1975). Some quantifiers for universal scope. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 37–58). New York, NY: Academic Press.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove, England: Erlbaum.

Johnson-Laird, P. N., & Byrne, R. M. J. (1993). Mental models or formal rules? *Behavioral and Brain Sciences*, *16*, 368–376.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, *99*, 418–439.

Lakoff, G. (1970). Linguistics and natural logic. *Synthese*, *22*, 151–271.

Lea, R. B. (1995). Online evidence for elaborative logical inference in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 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 processing. *Journal of Memory and Language*, *29*, 361–387.

Mandler, G. (2007). *A history of modern experimental psychology*. Cambridge, MA: MIT Press.

Oaksford, M., & Chater, N. (Eds.). (2010). *Cognition and conditionals: Probability and logic in human thinking*. Oxford, England: Oxford University Press.

O'Brien, D. P. (1995). Finding logic in human reasoning requires looking in the right places. In S. E. Newstead & J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honor of Peter Wason* (pp. 189–216). Hove, England: Erlbaum.

O'Brien, D. P. (2004). Mental-logic theory: What it proposes, and reasons to take this proposal seriously. In J. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 205–233). New York, NY: Cambridge University Press.

O'Brien, D. P., Braine, M. D. S., & Yang, Y. (1994). Propositional reasoning by model: Simple to refute in principle and in practice. *Psychological Review*, *101*, 711–724.

O'Brien, D. P., & Li, S. (2013). Mental logic theory: A paradigmatic case for empirical research on the language of thought and inferential role semantics. *Journal of Foreign Languages*, *36*, 27–41.

O'Brien, D., & Manfrinati, A. (2010). The mental logic theory of conditional propositions. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 39–54). Oxford, England: Oxford University Press.

O'Brien, D. P., Roazzi, A., Athias, R., & Brandão, M. C. (2007). What sorts of reasoning modules have been provided by evolution? Some experiments conducted among Tukano speakers in Brazilian Amazônia concerning reasoning about conditional propositions and about conditional probabilities. In M. Roberts (Ed.), *Integrating the mind: Domain general vs. domain specific processes in higher cognition* (pp. 59–81). Hove, England: Psychology Press.

O'Brien, D. P., Roazzi, A., Dias, M. G., Cantor, J. B., & Brooks, P. J. (2004). Violations, lies, broken promises, and just plain mistakes: The pragmatics of counterexamples, logical semantics, and the evaluation of conditional assertions, regulation, and promises in variants of Wason's selection task. In K. Manktelow & M. C. Chung (Eds.), *Psychology of reasoning: Theoretical and historical perspectives* (pp. 95–126). Hove, England: Psychology Press.

O'Brien, D. P., Roazzi, A., Dias, M. da G. B. B., & Soskova, J. (2007). Prevendo inferências lógico-predicativas no processamento de texto [Predicting predicate-logic inferences in text processing]. *Revista Interamericana de Psicología*, *41*, 119–128.

Over, D., & Evans, J. St. B. T. (2003). The probability of conditionals: The psychological evidence. *Mind & Language*, *18*, 340–358.

Passmore, J. (1994). *A hundred years of philosophy*. Harmondsworth, England: Penguin.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Wittgenstein, L. (1958). *Philosophical investigations*. Oxford, England: Basil Blackwell.

Yang, Y., Braine, M. D. S., & O'Brien, D. P. (1998). Some empirical justification of the mental predicate logic model. In M. D. S. Braine & D. P. O'Brien (Eds.), *Mental logic* (pp. 333–366). Mahwah, NJ: Erlbaum.

This is a section of doi:10.7551/mitpress/11252.001.0001

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

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