2 Logical Expressivism and the Open Structure of Reasons

Robert B. Brandom

2.1 A Pragmatic Definition of Implication and Incompatibility

Chapter One introduced a pragmatic metavocabulary: a vocabulary for specifying discursive practices and the basic roles performances can play in such practices. Specifically *discursive* practices are distinguished as those in which some performances are practically accorded the pragmatic significance of *claimings*. The idea is that discursive practices can be picked out as practices of *explicitly expressing* doxastic practical attitudes. They are practices that permit one to *say that* things are thus-and-so. Declarative *sentences* are identified as linguistic expressions whose free-standing (unembedded) utterance has the default significance of claimings.

Claimings are understood as essentially *bivalent* speech acts. They take two complementary forms: *assertions* and *denials*. Assertions express practical attitudes of *acceptance* (of what is expressed by declarative sentences) and denials express practical attitudes of *rejection* (of what is expressed by declarative sentences). One might, but need not, in addition understand the attitude of acceptance as *taking-true* and rejection as *taking-false*. However on the pragmatics-first order of explanation adopted here, the pragmatic attitudes of acceptance and rejection, and the speech acts of assertion and denial that express them, are conceptually and explanatorily prior to the concepts of <u>truth</u> and <u>falsity</u>.

A further overarching commitment of the present approach is a kind of *pragmatic linguistic rationalism*. It affirms the *critical rational* character of basic discursive practices. This is the claim that it is essential to the pragmatic significance of claimings as such that they are subject to rational *challenge* by some other claimings, and in need of rational *defense* by yet different claimings. Challenging and defending can be thought of as analogous to illocutionary forces claimings can have. Challenging is making claims that serve as *reasons against* the challenged claim. Defending (justifying) is making claims that serve as *reasons for* the challenged claim. The idea is that performances can have the pragmatic significance of claimings only in the context of practices of giving and asking for reasons. Claimings are what both can be given as reasons and for which reasons can be asked—what can both serve as and stand in need of reasons.

A *two-sorted deontic normative* pragmatic metavocabulary is apt for specifying the critical rational character of claimings. Claiming is understood as undertaking a distinctive kind of *commitment*. Commitments of this kind, *doxastic* commitments, come with a sort of default *entitlement*. The claimer is entitled to the doxastic commitment undertaken by claiming until and unless the claim is successfully challenged by claims expressing reasons against it. That entitlement can then be vindicated or regained by successfully defending it by claims expressing reasons for it.

In order to be intelligible as playing the role of claimables in critical rational discursive practices, what is expressed by declarative sentences must accordingly be understood as standing to what is expressed by other declarative sentences in two kinds of *reason relations*. Reasons *for* are claimables that *imply* the claimable they stand to as reasons for. Reasons *against* are claimables that are *incompatible with* the claimable they stand to as reasons against. The bivalence of doxastic attitudes of acceptance and rejection, manifested in the bivalence of speech acts of assertion and denial, together with the essential critical rational articulation of claimings as such together entail the *dyadic structure of reason relations*: their division into relations of *implication* and *incompatibility*. They may be thought of as relations of rational *inclusion* and *exclusion*. *Semantic inferentialism* is the thesis that the claimables expressed by the declarative sentences used to assert and deny are to be understood in terms of the roles those sentences play with respect to reason relations of implication and incompatibility. This is what we shall mean by "conceptual roles" and "implicational roles." Chapter Five will introduce what we claim is the intrinsic metavocabulary of conceptual roles: implication-space semantics.

The two-sorted deontic normative pragmatic metavocabulary is expressively sufficient to permit the *bilateral* definition of reason relations. A set of sentences Γ *implies* a sentence *A* ($\Gamma \mid\!\sim A$) if and only if *commitment* to *accept* all of Γ *precludes entitlement* to *reject A*. A set of sentences Γ is *incompatible* with a sentence *A* if and only if *commitment* to *accept* all of Γ *precludes entitlement* to *accept* A. If Γ *implies A* we can say that commitment to accept all of Γ *implicitly* commits one to *accept A*, equating preclusion from entitlement to reject with implicit commitment to accept. If Γ is *incompatible with A* we can say that commitment to accept all of Γ *implicitly* commits one to *reject A*. *Inferring* is *explicitly* acknowledging *implicit* commitments: actually accepting something that one's other commitments preclude one from entitlement to reject, or actually rejecting something that one's other commitments preclude one from entitlement to accept.

2.2 Reasons and Logic: Logicism about Reasons

So far we have articulated a version of a *pragmatic discursive rationalist* view that puts reasons and reason relations of consequence and incompatibility at the center of an account of what distinguishes specifically *discursive* practices. We have also gestured at a *semantic inferentialist* order of explanation that seeks to understand the conceptual contents declarative sentences express in terms of the role those sentences play in such reason relations—though this will not be a central topic of this work until Chapter Five. This chapter and the next are devoted not to pragmatics or semantics, but rather to *logic*. They are animated by the conviction that an axial issue in the philosophy of logic is what might be called "the reasons question about logic." How should we understand the relations between logic and reasons?

There is a long tradition that takes logic to be a study of reasons. The reasons question about logic asks for further specification of the distinctive *kind* of inquiry into reasons logic is: its characteristic methods, goals, and presuppositions. Frege's 1879 *Begriffsschrift*—the founding document of modern logic—offers a resounding and sophisticated response to this challenge, of the expressivist sort that we will be recommending and pursuing. But his mature work, in the 1890's, shifts focus, from reasons to truth. Michael Dummett says of this shift:

> ... in this respect (and in this respect alone) Frege's new approach to logic was retrograde. He characterized logic by saying that, while all sciences have truth as their goal, in logic truth is not merely the goal, but the object of study. The traditional answer to the question what is the subject-matter of logic is, however, that it is, not truth, but inference, or, more properly, the relation of logical consequence. This was the received opinion all through the doldrums of logic, until the subject was revitalized by Frege; and it is, surely, the correct view.
>
> (Dummett, 1973, 432)

And again:

> It remains that the representation of logic as concerned with a characteristic of sentences, truth, rather than of transitions from sentences to sentences, had highly deleterious effects both in logic and in philosophy. In philosophy it led to a concentration on logical truth and its generalization, analytic truth, as the problematic notions, rather than

> on the notion of a statement's being a deductive consequence of other statements, and hence to solutions involving a distinction between two supposedly utterly different kinds of truth, analytic truth and contingent truth, which would have appeared preposterous and irrelevant if the central problem had from the start been taken to be that of the character of the relation of deductive consequence.
>
> (Dummett, 1973, 433)

Dummett is here acknowledging the centrality of the reasons question to the issue of what logic is. It is, he says, the study of a certain kind of *consequence* relation: "logical" consequence or "deductive" consequence. One way of sharpening the reasons question then is to ask how we should understand the relations between *logical* consequence or implication, and consequence or implication in *general*. Though he does not do so in these passages, presumably Dummett would acknowledge that logic also studies the notion of inconsistency, which is specifically *logical* incompatibility. And once again, we can ask how this special logical kind of incompatibility is related to the incompatibility of claimable conceptual contents in general.

The reasons question invites us to consider two contrary orders of explanation: understanding and explaining *reasons* in terms of *logical* reasons, and understanding *logical* reasons in terms of reasons more generally.[1] Frege pursued a logicist reconstruction of central mathematical concepts, explaining the reason relations they stand in to one another entirely in terms of specifically *logical* reason relations. His striking success (in spite of some attendant sophisticated and telling failures) was a testament to the expressive power of his new logic. It inspired a characteristic, axial project at the core of what was to become analytic philosophy: to reconstruct the reason relations that articulate conceptual contents in general using only *logically valid* relations of consequence and inconsistency. That is to do for important concepts outside mathematics what Frege had sought to do for central concepts within mathematics.

The master idea is that "*good* reason" just means "*logically* good reason." Where this is not obvious, it is the philosopher's task to unearth the hidden logical articulation of apparently nonlogical concepts, so as to be able to reconstruct, in purely logical terms, the relations of consequence and incompatibility they stand in to other concepts. The paradigm epitomizing this enterprise was Russell's 1905 essay "On Denoting." It discerns a logical deep structure underneath the definite descriptions of ordinary nonlogical descriptive vocabularies—expressions such as "the author of *Waverly*" and "the present King of France." Russell then shows how the consequences and incompatibilities of different kinds of sentences in which definite descriptions occur can be derived *logically*, as valid implications and provable inconsistencies, from their suggested reformulations in logical terms.

The project launched by what Russell made of Frege is *logicism about reasons*, extended and generalized from its original form, which was addressed specifically to mathematical reasoning, to include the reason relations articulating *all* legitimate conceptual contents. Wittgenstein's *Tractatus* distills this logicism to its pure, crystalline form. In its subsequent empiricist form, logicism is the animating idea of the program of Carnap's Vienna Circle. Nonlogical content is provided to empirical concepts by their relations to perceptually occasioned noninferential observation reports, but the implications and exclusions of one another by those empirical concepts are to be shown to be matters exclusively of *logical* consequence and inconsistency, by suitable reconstructions or excavations of the logical form those concepts give to that raw nonlogical, nonconceptual, empirical content. According to this way of thinking about the relations between logic and reasons, logic is not just the *study* of reasons, it is the *science* of reasons. Logic sets the standards deciding what is a genuine reason for and against what, no matter the nonlogical subject matter being reasoned about.

Thought of this way, the logicist answer to the reasons question is recognizably a successor conception to what is arguably a founding idea of the Western philosophical tradition. The Greeks were puzzled and fascinated by the peculiar normative force of the better reason. How was this unique kind of specifically *rational* compellingness to be distinguished from or within the genus of utterances that merely reliably dispose their audiences to assent, without providing justifying reasons—the functional characterization of Sophists' techniques? Plato found his answer in the geometry that Euclid would go on to codify. For there, not acknowledging as compelling an argument in the form of a proof just shows that one did not understand it. In geometrical argumentation reasons display their force unhindered and unobscured. The philosophical enterprise is coeval with the aspiration to understand *all* reasons on this model.

The path from this mathematical model of reasons in geometrical form to a logicist model is long, and distinguished. While accepting Galileo's claim that the book of nature is written in the language of mathematics, Descartes generalized philosophically from his own analytic geometry and understood reasoning about the geometrical world of extension in terms of the manipulation of algebraic symbols. Spinoza hews to the geometrical model of reasons, subtitling his book on Descartes "*More Geometrico Demonstrata*," and his *Ethics* "*Ordine Geometrico Demonstrata*." Leibniz could envisage a future logic as a *calculus ratiocinator* only because he was looking beyond the syllogistic logic of the schools towards an ideal, specifically mathematical, logic in the form of an algebra. Nor was this line of thought confined to rationalists. Hobbes expressed his endorsement of the mathematical model of reasoning in general in the slogan "Reasoning is but reckoning" (meaning arithmetic calculation). It was not until Frege's *Grundgesetze* and Russell and Whitehead's *Principia Mathematica* worked out in concrete detail projects for reconstructing substantial mathematical reasoning in terms of their new, expressively powerful logics that the general logicist program of understanding all theoretical reason relations as disguised logical reason relations could give shape and impetus to the emergent, ascendant school of analytic philosophy.

The generalization of logicism about specifically mathematical reasoning to logicism about reasons *tout court* was a bold and original philosophical thesis. Looking back over the results of more than a century's efforts in pursuing its promises and working this claim out in detail raises the suspicion that, as Hempel says of the Vienna Circle's attempts to formulate an empiricist criterion of meaningfulness, it

> … has come to play the role of the treasure hunt in the tale of the old winegrower who on his death-bed enjoins his sons to dig for a treasure hidden in the family vineyard. In untiring search, his sons turn over the soil and thus stimulate the growth of the vines: the rich harvest they reap proves to be the true and only treasure in the vineyard.
>
> (Hempel, 1963, 707)

We will see that the general logicist model of reasons faces deep difficulties of principle, rooted in the most general features of the structure of reason relations of consequence and incompatibility. Logicism about reasons entails that reason relations in general have (or can be reduced to logical relations that have) the topological *closure* structure characteristic of logical reason relations. If that *structural* logicist claim is not correct, then logicism about reasons cannot be right either.

2.3 Reasons and Logic: Rational Expressivism about Logic

Before turning to that structural argument against logic-first explanations of the relations between logic and reasons, it is worth considering for comparison a converse approach that pursues a complementary, reasons-first order of explanation. Logical *expressivism* is the view that logic is a *language* of reasons rather than a *science* of reasons. It is not a theory that determines or explains what is a reason for and against what. The distinctive task of logical vocabulary is not to *prove* that one set of claims implies or is incompatible with another, or to *derive* what is a reason for and against what, but rather to make it possible to *say* what is a reason for and against what. In this sense, logic is not a *canon*, but an *organon* of reason:

not a set of principles determining or explaining reason relations but a set of expressive tools for making reason relations explicit. Making implications and incompatibilities explicit is codifying them in the sentential form of claimables, that is, as the sort of things that can themselves be given as reasons and for which reasons can be asked. The point of adding logical vocabulary to a base vocabulary is to make it possible to talk about the reason relations of that base vocabulary—and so to reason about those reason relations, in the sense of making claims about them, and challenging and defending those claims.

Expressivism is the philosophical understanding of the characteristic, defining task of logic that Frege introduces in his seminal *Begriffsschrift* of 1879. To introduce his project he says:

> ... there are two ways in which the content of two judgments may differ; it may, or it may not, be the case that all inferences that can be drawn from the first judgment when combined with certain other ones can always also be drawn from the second when combined with the same other judgments. The two propositions 'the Greeks defeated the Persians at Plataea' and 'the Persians were defeated by the Greeks at Plataea' differ in the former way; even if a slight difference of sense is discernible, the agreement in sense is preponderant. Now I call that part of the content that is the same in both the conceptual content. Only this has significance for our symbolic language [*Begriffsschrift*] ... In my formalized language [BGS] ... only that part of judgments which affects the possible inferences is taken into consideration. Whatever is needed for a correct ['richtig', usually misleadingly translated as 'valid'] inference is fully expressed; what is not needed is ... not.
>
> (Frege, 1998, sec 3)

His logical notation is conceived as a tool for expressing the inferences sentences are involved in. And he introduces a special notion of content, *conceptual* content (*begrifflicher Inhalt*), which he identifies with the role sentences play in inferences (*Schlüsse*)—that is, their role in what we have been calling "reason relations." Expressing inferential relations is a way of making conceptual contents explicit. This is a semantic inferentialist thesis. It is why he calls his logic a "concept vocabulary," a way of expressing concepts in writing: *Begriffsschrift*. (His two-dimensional logical notation can *only* be written, not spoken.) Describing this project afterwards, Frege says:

> Right from the start I had in mind the expression of a content ... But the content is to be rendered more exactly than is done by verbal language ... Speech often only indicates by inessential marks or by imagery what

> a concept-script should spell out in full.
> (Boole's logical Calculus and the Concept-script, in Frege, 1979, 12-13)

This is logical expressivism: logic provides tools for characterizing reason relations, specifying what follows from what and what is incompatible with what.

It is important that as examples of sentences whose conceptual content (inferential roles) his logical vocabulary serves to explicate he offers *nonlogical* sentences, in the sense of sentences that contain no specifically logical vocabulary. The conceptual contents he aims to "spell out in full" to begin with are mathematical concepts concerning numbers. But he is clear that his expressive ambitions extend much further:

> It seems to me to be easier still to extend the domain of this formula language to include geometry. We would only have to add a few signs for the intuitive relations that occur there ... The transition to the pure theory of motion and then to mechanics and physics could follow at this point.
> (Frege, Begriffsschrift Preface, in van Heijenoort, 1967, 7)

He thinks that sentences with these nonlogical conceptual contents, for instance, those of physics, already stand in reason relations of implication and incompatibility to one another before logic comes on the scene. The task of logic is to "render those contents more exactly" by making explicit those nonlogical reason relations. Since he is a semantic inferentialist, understanding the conceptual contents of sentences in terms of their roles in reason relations, Frege thinks that by making it possible to specify reason relations clearly and precisely, logical vocabulary thereby enables perspicuous representation of the conceptual contents expressed by sentences formulated using nonlogical vocabulary.

According to such an expressivist view, the central, defining use of logical vocabulary is its use in conjunction with some nonlogical base vocabulary, whose reason relations it makes explicit. Frege explicitly mentions as examples of target base vocabularies, to begin with the vocabularies of arithmetic and geometry in mathematics, and the vocabularies of kinematics and dynamics in physics. Those base vocabularies are not just unstructured sets of lexical items. They are construed as coming with their own native reason relations of implication and incompatibility. These are what Sellars calls "material" inferential relations. He uses that term to emphasize that they are not relations of *logical* consequence and inconsistency. They articulate the conceptual contents of the *non*logical concepts that are deployed in the base vocabulary—as for instance the goodness of implications whose premises are of the form "*A* is to the East of *B*," and whose conclusions are of the form "*B* is to the West of *A*," is part of what makes the nonlogical terms "East" and "West" mean what they mean, express the conceptual contents that they do. Newtonian physics determines what implications relate claims about the mass and acceleration of a body to claims about the forces acting on it. The expressivist idea is that logical vocabulary can then be added to such a base vocabulary, to codify its material reason relations in perspicuous form.

Of course, for the addition of logical vocabulary to a base vocabulary to make it possible to express the reason relations of the base vocabulary, the reason relations in which sentences of the logically extended vocabulary stand to one another must also be determined. Doing that is settling how the implications and incompatibilities of logically complex sentences formed by applying logical operators to sentences of the base vocabulary systematically inherit their reason relations from those of the base vocabulary. It turns out that that can be done, for instance in Gentzen-style sequent calculi, in a purely logical metavocabulary. That is, one can specify the reason relations that hold among sentences that contain no logical terms in such a way that it is settled thereby how the reason relations that hold between arbitrary sentences of *any* logically extended base vocabulary are determined by the material reason relations native to the base vocabulary. This is a remarkable fact, and a substantial achievement. (In Chapter Three we will look in some detail at how this works for our specific candidate for an expressively powerful logic.) It is this fact about the specifiability of purely logical reason relations, together with the fact that logical vocabulary so articulated can then be used to codify the reason relations of any base vocabulary whatsoever, that makes it tempting to conjecture that all material reason relations are made good by implications and incompatibilities that hold in virtue of logic alone—as the logicist about reasons does.

The logicist takes the home language-game of logical locutions—the sort of use that defines them as *logical* locutions—to be *pure* logic. On the side of the lexicon, pure logic is concerned with sentences that contain *only* logical vocabulary. In the sentential case, with which we are exclusively concerned in this work, this means sentential connectives, such as conditionals, negation, conjunction, and disjunction, together with schematic variables standing in for the sentences they are applied to, from which logically complex sentences are formed. And on the side of reason relations, the vocabulary of pure logic considers only relations of logical consequence and inconsistency that hold just in virtue of those logical connectives: for example the implication of A by $A \wedge B$ and the incompatibility (logical inconsistency) of A and $\neg A$. As is clearest in its explicit Tractarian form, the logicist aspiration is then to explain all genuine reason relations among sentences containing nonlogical terms by understanding some of them as covertly having the form of logically complex sentences formed from other nonlogical expressions by applying logical connectives to them and deriving the good implications and incompatibilities that relate them as logically valid implications and logical inconsistencies.

Expressivism claims that if one thinks only about the relations between logical vocabulary and purely logical reasons, one will miss the most important relations between logic and reasons. For the expressivist, by contrast to the logicist, pure logic is a by-product of what is actually the characteristic use that marks some locutions as *logical* locutions. That characteristic, defining use is the application of those locutions to extend *non*logical base vocabularies, by adding logically complex sentences formed from the sentences of the base vocabulary by applying logical sentential connectives recursively to them. The relation between the base vocabulary and its logical extension is two-fold. The logically extended vocabulary is *elaborated from* the base vocabulary, and it is *explicative of* it. We abbreviate this two-fold functional specification of the expressive role characteristic of logical vocabulary by saying that it is "LX" for its base vocabulary: *el*aborated from and *ex*plicative of it.[2] The central idea of the "elaboration" side of this formula is that both the lexicon of sentences of the logically extended vocabulary and the reason relations of implication and incompatibility that articulate the conceptual roles of those sentences are uniquely and completely determined by and computable from the lexicon and reason relations of the base vocabulary that is logically extended. The central idea of the "explication" side of this formula is that the extension of the base vocabulary by the addition of logical vocabulary makes it possible to *say*, in the logically extended vocabulary, what implication and incompatibility relations hold not only in the base, but also in its logical extension.

These two sides of the functional characterization of the expressive role that defines logical vocabulary are not independent of one another. The explicative expressive function imposes an important criterion of adequacy on the rules for elaborating relations of implication and incompatibility for the logically extended vocabulary from the relations of implication and incompatibility that govern the material base vocabulary. The rules for introducing logical connectives—paradigmatically those that ensure that the conditional codifies implication and negation codifies incompatibility in accordance with the principles formulated above—must extend the relations of implication and incompatibility that they elaborate *conservatively*. That is, all the inferential relations among nonlogical vocabulary must be preserved, and no new implications or incompatibilities involving only nonlogical vocabulary must be introduced.

Why? Because introducing vocabulary to *express* those reason relations should not *change* them. It should *just* make it possible to express *explicitly* in the (logically extended) object language the reason relations that *implicitly* govern the material base vocabulary (which we theorists express in a proof-theoretic *meta*language of sequents). We will see below that there are other notions of explication where it is not appropriate to impose a corresponding conservativeness condition. But conservativeness is required for the sort of explication that governs the introduction of specifically *logical* vocabulary.

A particularly vivid example of the trouble one can get into if one does not impose this condition is provided by the connective "tonk" introduced by Arthur Prior as part of an argument against natural deduction calculi without restrictions on the rules. He pointed out that the effect of using "tonk" with the introduction rule of classical disjunction, which includes:

$$\frac{\text{p}}{\text{p tonk q}}$$

and the elimination rule of classical conjunction, which includes:

$$\frac{\text{p tonk q}}{\text{q}}$$

is to license the implication:

$$\cfrac{\cfrac{\text{p}}{\text{p tonk q}}}{\text{q}}$$

He called that a "runabout inference ticket," since it licenses the implication from arbitrary logically atomic *p* to arbitrary logically atomic *q*. Nuel Belnap diagnosed the trouble as a violation of conservativeness.[3] Introducing new logical vocabulary should not create any new implications involving only old vocabulary. Ignoring that constraint courts the risk of "tonking up" one's implication relation: trivializing it. The expressivist's requirement that logical vocabulary be explicative of the reason relations of the base vocabulary it extends offers a principled rationale for the requirement of conservativeness that Belnap imposed more or less *ad hoc*.

Although conservativeness is a criterion of adequacy (and so a necessary condition) for any vocabulary playing the distinctive explicative expressive role being recommended as demarcating specifically logical vocabulary, it is not appropriate to require it for nonlogical vocabulary in general. For it is characteristic of ordinary empirical descriptive vocabulary that its use involves endorsing nontrivial material implications relating its circumstances of appropriate application to its appropriate consequences of application. The nonlogical content of concepts such as <u>copper</u>, <u>fever</u>, and <u>cruel</u> incorporate many such substantive material implications (and incompatibilities) that need not be redundant relative to other concepts antecedently available. This important dimension of content is ignored and made invisible by understanding conceptual content in terms of truth conditions. For that is the idea of conditions that are both individually necessary and jointly sufficient—the idea of circumstances and consequences of application that are guaranteed to coincide, so that no substantive inferential commitment is involved in the transition between them.

The covert ideology behind the notion of truth conditions is accordingly a kind of logicism. It is the idea that the reason relations that govern *all* vocabulary should not just be *expressible* by logical vocabulary (vocabulary governed by *logical* reason relations), but that those reason relations must be in the end *reducible to* logical reason relations. Since logical reason relations must be conservative over the rest of the language, so must reason relations in general. That is the origin of the requirement of the coincidence of circumstances and consequences of application in the form of conditions that are both necessary and sufficient. But that squeezes out nonlogical content, as articulated by *materially* good relations of implication and incompatibility: those that hold in virtue of the contents of *non*logical concepts. Here logicism about the relations between logic and reason relations misleads us about the latter, and does damage to our semantic theory (according to even a very weak form of semantic inferentialism).[4]

So the first half of the "LX" characterization of the expressive role characteristic of logical vocabulary is that the reason relations of implication and incompatibility that govern logically complex sentences must be conservatively elaborated from the reason relations governing the logically atomic sentences of some base vocabulary that is extended by the addition of those logically complex compounds of the atomic sentences. We use "vocabulary" here in a technical sense, as meaning a relational structure comprising a *lexicon*, which is a set of sentences, together with reason relations of implication and incompatibility defined on that lexicon. Syntactic formation rules must compute the lexicon of the logically extended vocabulary from that of the base vocabulary by the application of connective rules—which say, for instance, that if A and B are sentences in the extended lexicon, then so is $A \to B$. And an elaboration function must be specified that shows how to compute the reason relations governing sentences of the extended lexicon conservatively from reason relations governing sentences of the base vocabulary.

It is much less straightforward to say exactly what is meant by the portion of the LX-ness condition that requires the logically extended vocabulary that is elaborated from a base vocabulary to be *explicative of* the reason relations of the base vocabulary. We will have a lot more to say about this issue in the next chapter, when we introduce our preferred universally LX logic, NMMS. The underlying thought is simple enough. For the expressivist, the paradigmatic logical connectives are the conditional and negation. The characteristic and defining expressive function of *conditionals* is to make explicit *implication* relations, and the characteristic and defining expressive function of *negation* is to make explicit *incompatibility* relations. Those connectives do that best if they satisfy these two conditions:

- Deduction-Detachment (DD) Condition on Conditionals:
  $\Gamma \mathrel{|\!\sim} A \to B$ if and only if $\Gamma, A \mathrel{|\!\sim} B$.
- Incoherence-Incompatibility (II) Condition on Negation:
  $\Gamma \mathrel{|\!\sim} \neg A$ if and only if $\Gamma \# A$.

The first condition is inspired by Frank Ramsey's understanding of the expressive role of conditionals, and could with equal justice be called the "Dual Ramsey" Condition. It says that a premise-set implies a conditional just in case the result of adding the antecedent of that conditional to it is a premise-set that implies the consequent of the conditional. In this way, facts about implications are expressed in conditional sentences. Deduction-Detachment thereby offers one clear sense in which the conditional $A \to B$ *says that* A implies B. For if the underlying nonlogical base vocabulary contains the implication

It is raining $\mathrel{|\!\sim}$ The streets will be wet

(which is materially a good implication in part because of the meanings of the non-logical terms "raining" and "wet"), and it does not contain the implication

The coin in my hand is copper $\mathrel{|\!\sim}$ The coin in my hand would melt at 100°C.

then in the vocabulary that has been logically elaborated from that base, the conditional

It is raining $\to$ The streets will be wet

is provable or derivable from the first sequent, and there is no way so to become entitled to the conditional

The coin in my hand is copper $\to$ The coin in my hand would melt at 100°C.

in the absence of the second sequent. Neither one holds in virtue of logic alone, as does the more complex conditional

> ((It is raining → The streets will be wet) ∧ It is raining) → The streets will be wet.

which follows from Deduction-Detachment alone (in the context of Containment). The expressivist idea is that it is the fact that the nonlogical conditional can be justified (here, in a sequent calculus with sequents from the nonlogical, material base as leaves of the proof tree), and what it can in turn be used to justify that is the key to understanding the expressive role of conditionals. The fact that the more complex conditional holds just in virtue of the conceptual content of the conditional merely reflects that underlying expressive role.

The Incoherence-Incompatibility condition on negation says that a premise-set Γ implies a negated sentence just in case the sentence negated is incompatible with Γ. It offers one clear sense in which the negation ¬*A says that A* is incompatible with a premise-set—namely, any premise-set that implies that negation. For it codifies an understanding of the negation of *A* as its minimal incompatible. ¬*A* is implied by everything that is materially incompatible with *A*. The Aristotelian contradictory of *A* is what is implied by all the Aristotelian contraries of *A*. So not-red is implied by green, blue, and so on. In the logically extended vocabulary, one can say that "the swatch is green" is incompatible with "the swatch is red" by asserting the conditional with negated consequent

> The swatch is green → ¬(The swatch is red).

Logicists want to explain material incompatibility (Aristotelian contrariety) in terms of logical negation—and so, inconsistency (Aristotelian contradiction). Expressivists seek to exploit these same conceptual connections in the other direction. (In the useful terms current in the philosophy of logic, the order of explanation that introduces the concept of logical negation on the basis of a prior concept of material incompatibility is understanding negation on the "Australian plan," by contrast to the contrary "American plan" (Berto and Restall, 2019).)[5]

Logical expressivists focus on the directly *expressive* sentential logical connectives: conditionals and negation. These contrast with merely *aggregative* logical connectives, paradigmatically conjunction and disjunction. The function of these auxiliary connectives is to make explicit what is indicated by commas on the two sides of the turnstile in a sequent-calculus formulation of reason relations. We sometimes refer derisively to the merely aggregative connectives as "Boolean helper-monkeys."

Chapter Three shows how the expressivist aspirations sketched here in terms of the LX-ness of logic for arbitrary base vocabularies can be made more definite, and then how those aspirations can be realized in the familiar form of a multisuccedent sequent calculus. We also show there how the logic we endorse can be enriched with modal operators whose expressive role is to make explicit features of the local and global *structure* of the reason relations that are in turn made explicit by ordinary, nonmodal sentential connectives.

Thinking about the relations between reasons and logic according to the expressivist model being recommended here offers a distinctive perspective on traditional philosophical notions of rational consciousness and self-consciousness thematized most prominently by classical German idealists. Adopting practical discursive attitudes of accepting and rejecting claimables, which stand to one another in reason relations of implication and incompatibility, defines a distinctive sense in which creatures can qualify as *conscious*. This is consciousness in the sense of sapience or apperception, rather than sentience or perception. That sort of conceptually articulated awareness is not just a matter of being awake and sensorily responsive—able, say, to feel pain and distinguish red things from green ones. It consists rather in taking it *that* things are thus-and-so, in having propositionally contentful attitudes expressible by what count as declarative sentences just in virtue of expressing contents that can stand to one another in material reason relations of implication and incompatibility—to serve as and stand in need of reasons. Apperceptive consciousness in this sense is accordingly to be understood as *rational* consciousness: consciousness articulated by reasons.

Understanding the relation of logic to those reason relations as the logical expressivist does then presents logical vocabulary as the organ of a distinctive kind of rational *self*-consciousness. For it enables the making explicit in claimable form of the reason relations that articulate rational, apperceptive consciousness. Chapter Three shows how this sense of specifically *logical* rational self-consciousness as logical explicitation can be made precise in the form of a powerful expressive completeness theorem (tied to what we call "explicitation by sequents") relating what can be said in the logical extension of a base vocabulary to the reason relations of that base vocabulary.

The two complementary approaches to the reasons question about logic, logicism about reasons and rational expressivism about logic, agree in one feature of their ambition: its generality. Logicism seeks to exhibit *all* (doxastic, nonpractical) reasons as covertly *logical* reasons—good, if and insofar as they are good reasons, in virtue of being reducible to logically valid implications (in the case of reasons-for) and logical inconsistencies (in the case of reasons-against). Analogously, the rational expressivists' ideal logic would be LX for *every* base vocabulary. The connective rules defining its logical vocabulary would permit it conservatively to extend, elaborate its reason relations from, and articulate the conceptual contents of, logically complex sentences that express the reason relations of *any* base vocabulary whatsoever. Expressivists can recognize vocabularies with more limited expressive power as genuinely logical—for instance, logics that can codify only mathematical implications, or (more to the point for our larger argument here) only structurally closed, monotonic and transitive reason relations. Insofar as they are LX for some restricted range of base vocabularies, they can still qualify as proper logics for the logical expressivist. This is the sense in which there can be logics of colors, or sortals, or quantum phenomena—just insofar as the vocabulary in question can be elaborated from and is explicative of color vocabularies, sortal terms, or the vocabulary of quantum mechanics. But the expressivist ideal is expressive *universality*: a logic that is LX for any vocabulary whatsoever, no matter what its lexicon or reason relations might be. The ultimate measure of a logic—the normative standard for assessments of its success *as* a logic—is its expressive power, the range of reason relations it can make explicit. Perhaps universal LX-ness is unobtainable, practically, or in principle. But anything less counts as expressive impoverishment.

From the point of view of rational expressivism about logic, this ideal shows up as what is right about traditions in the philosophy of logic that pick out purely logical reason relations as those that are topic neutral and (so) maximally general. Logical reasoning, the thought goes, applies to any and every subject matter, by contrast to content-specific reason relations that hold only for some topics. One way of articulating this thought originates with Bolzano, is exploited by Frege, and is given voice by Quine. It is a methodology of noting invariance of reason relations under substitution.[6] An implication or incompatibility holds in virtue of the logical form of the sentences involved just in case it is robust under the (uniform) substitution of nonlogical for nonlogical vocabulary. Considering reason relations that are invariant under arbitrary substitution of nonlogical for nonlogical vocabulary is a clear and precise way of operationalizing the idea that the reason relations being assessed hold independently of the nonlogical subject-matter of the sentences involved.

This is an idea logical expressivists can wholly endorse. For notice that it presupposes a notion of the goodness of *material* (nonlogical) implications and incompatibilities. It picks out some reason relations as good in virtue of their purely logical form in case two conditions hold:

i) They are good, and

ii) All of their substitutional variants, resulting from replacement of nonlogical by nonlogical vocabulary, are also good.

(ii) is the condition that the implications and incompatibilities be good under arbitrary substitutions *salva consequentia*. Both conditions appeal to a notion of the goodness of an implication or incompatibility relation that is *not* a matter of their *logical* goodness (that is what is being defined), and must be intelligible in advance of the notion of formal logically valid implications and inconsistencies, so as to be available for use in the definition. Of course, for the expressivist about logic, the procedure only works if one has available a vocabulary that includes logically complex sentences formed from the sentences of some nonlogical base vocabulary—that is, vocabularies that result from the logical elaboration of those base vocabularies. The dependence of this way of understanding and justifying the universality of logic on prior material relations of implication and incompatibility puts it in substantial tension with the logicist's aspiration to understand those relations in terms of implications and inconsistencies that hold in virtue of logic alone.

It is worth noticing in passing that the Bolzano-Frege-Quine method of noting invariance under substitution is really a way of thinking about the concept of form generally, rather than a way of thinking about *logic* and *logical* form specifically. For it does not at all depend on the vocabulary that is kept invariant being logical vocabulary. Any vocabulary at all can be substitutionally privileged for these purposes. If one looks at materially good implications that remain good under arbitrary substitution of non-*theo*logical for non-*theo*logical (or non-geological for non-geological, or non-culinary for non-culinary, non-nautical for non-nautical …) vocabulary, the result will be implications that hold in virtue of their theological (geological, culinary, nautical …) form alone. (These might relate terms such as "sacred" and "profane", "sin" and "peccable," and so on.) In particular, when the substitutional methodology is applied to logic as supplying the invariance-privileged vocabulary, it presupposes a distinction between logical and non-logical vocabulary. That methodology accordingly cannot be appealed to in demarcating logical vocabulary. Our expressivist proposal for that demarcational task is that logical vocabulary is vocabulary that is LX for some base vocabulary—in the ideal limit, *universally* LX, in being capable of being elaborated from and explicative of any and all base vocabularies.

2.4 The Structure of Reason Relations

The logical expressivist aspiration to universal explicative generality, in the sense of seeking a logic that can codify any and all reason relations, has one consequence that is of particular significance for the logical and semantic innovations that we will propose here. That is that an ideal logic would not build in any undue restrictions on the *structure* of the relations of implication and incompatibility that it can express. This observation points to a fundamental issue that is in general not accorded the centrality in contemporary philosophy of logic that we believe it deserves. There is a mismatch between the structure of specifically *logical* reason relations and the structure of reason relations in general.

For most of the last century it has been agreed (following Tarski (1936) and Gentzen (1934)) that logical consequence has the topological structure of a closure operator. That means that for any premise-set X (a set of sentences drawn from some lexicon), it makes sense to talk about *the* set $c(X)$ of its logical consequences. That consequence-set $c(X)$ is the *closure* of the premise-set X under logical consequence in the sense that it meets three structural conditions:

- The premises are included in the consequence set: $X \subseteq c(X)$. (Containment)
- Adding further premises does not lose any consequences: $c(X) \subseteq c(X \cup Y)$. (Monotonicity)
- The consequences of the consequences of any premise-set are already consequences of the original premise-set: $c(c(X)) = c(X)$. (Idempotence, entailing Transitivity)

The consequence relations of classical and intuitionist logics, and the modal logics built on them such as S4 and S5, as well as traditional multivalued logics such as Strong Kleene (K3) and its dual, Priest's Logic of Paradox (LP) are all structurally closed in this topological sense (defined in topology by the Kuratowski closure axioms, which inspired Tarski's metalogical application of them).

The implication relations that govern most of our practices of challenging and defending claims do not have this closure structure, though. From the point of view of traditional logic, ordinary material non- or prelogical reasoning (as well as its sophisticated institutionalized versions in law, medicine, and the special sciences) is in general *sub*structural. We will say that it has an *open* structure, by contrast to the *closed* structure of traditional logics. The substructurality of reasoning in general is most obvious in the failure of monotonicity. Almost always when we give reasons for or against our own assertions or those of others, those reasons are defeasible. That is, there are further considerations that, if admitted as auxiliary hypotheses or collateral premises, would infirm the implication or incompatibility connecting the proferred reasons to what they are presented as reasons for or against. We take such defeasibility for granted in casual justifications of or challenges to quotidian claims. If someone argues "Everyone is already in their offices, so we will be able to start the meeting on time," they know full well that a myriad of interfering events—last-minute telephone calls, bouts of forgetfulness, fire alarms ...—could infirm the implication. But more serious scholarly and scientific arguments are like this, too. The carbon 14 dating results provide good reasons for dating the White Sands footprints to 20,000 years ago—but not if the association of the footprints with the *Ruppia* seeds that were actually dated is spurious, or if the hard-water effect had incorporated more ancient carbon atoms into those seeds, and so on.[7] The reasons offered supporting or contesting medical diagnoses or clinical predictions are understood to be overrideable by all sorts of possible further evidence. The same is true of legal argumentation, where distinctive forms of defeasibility of reasoning are institutionalized in the form of appeals courts.

As was pointed out in the Introduction, two broad and important areas of philosophical thought about reasoning where the nonmonotonicity of implications is explicitly acknowledged, thematized, and theorized about are probabilistic reasoning and implications codified in subjunctive conditionals. An argument that justifies with high probability a certain conclusion about a population given one specification of the reference class from which the sample is drawn can turn into a much weaker argument, or even an argument with equally high probability for a contrary conclusion, if further information is added that narrows the reference class. Adding further premises can flip the valence of subjunctive reasoning:

> If I were to press this key, the program would start.
> If I were to press this key and Dan rewrote the code last night to require a different key, then the program would not start.
> If I were to press this key and Dan rewrote the code last night to require a different key, but Ulf noticed this morning and switched it back, then the program would start.
> But not if, in addition, the computer is unplugged.
> Unless it has a battery backup.
> But even then, if the battery has run down, the program won't start.
> ...

One might be tempted to treat all of these familiar phenomena as reflecting low pragmatic considerations of conversational convenience, rather than deep structural features of implications and incompatibilities. The thought would be that it is simply too laborious and tedious to put in all the qualifications required for fully stating a *real* reason—that is, an indefeasible, dispositive reason. So one just gestures in the direction of the complicated, extensive conjunction of evidentially significant premises or conditions that is implicitly being invoked when reasons are offered. But that is not right. For in the sorts of cases being considered, there typically *is* no totality of potentially relevant considerations, no full specifications of the further circumstances under which an implication could fail. Any finite list will be found to have left out some outré possibilities. The hungry lioness notices the nearby wounded gazelle, so she will chase it. But not if she is paralyzed, struck by lightning, suddenly encased in glass, reduced to microscopic size, teleported to Venus, and so on. And there is in general no nontrivial way to specify an infinite list that would be comprehensive. By "nontrivial" here is meant any way that is not equivalent to saying "The hungry lioness notices the nearby wounded gazelle, so she will chase it—except if for some reason she does not." If "all other things being equal" qualifications had the effect of ruling out all potentially defeating further premises, then such *ceteris paribus* clauses would be trivial in this objectionable way. In fact they should not be understood as somehow *removing* defeasibility and rendering implications monotonic. (A Latin phrase whose utterance can do *that* is properly called a "magic spell.") The real expressive function of *ceteris paribus* clauses is just to acknowledge explicitly that the implication being endorsed *is* defeasible. Such a concession is not a taking-back of the argument to which it is appended. Rather, it pragmatically marks the openness of a conversational pathway in which some of the open-ended set of possible defeaters are actually asserted, and so themselves become subject to critical rational challenge and defense.

Note that monotonicity is as well-defined for incompatibility as for implication. An incompatibility property is nonmonotonic if adding elements can make incompatible sets compatible. And monotonicity is equally implausible as a constraint on material incompatibility relations. Indeed, there is a general procedure for turning failures of monotonicity for material consequence relations into failure of monotonicity for material incompatibility relations. The nonmonotonicity of the implications codified in subjunctive conditionals is illustrated most forcefully by the possibility of "Sobel sequences," in which further information sequentially flips the valence of the implication. If I were to strike this dry, well-made match, it would light. But *not* if it is in a very strong magnetic field. Unless, additionally, it were in a Faraday cage, in which case it would light. But *not* if the room were evacuated of oxygen. And so on. But here we can see that reasons *against* a claim are as defeasible in principle as reasons *for* a claim. Striking the dry, well-made match and its lighting is incompatible with its being in a strong magnetic field. But striking the dry, well-made match and its lighting and its being in a Faraday cage *is* compatible with its being in a strong magnetic field. Negation operators introduced to codify material incompatibility relations will have to deal with the structural nonmonotonicity of incompatibility every bit as much as conditionals introduced to codify material consequence relations will have to deal with the structural nonmonotonicity of implication. That this observation about the open structure of both kinds of reason relation is of some importance becomes clear in light of the fact that prominent approaches to nonmonotonic logics (such as preferential models and default reasoning) help themselves to classical negation and specifically, the monotonic relation of inconsistency it supports (Kraus et al., 1990; Horty, 2007).[8]

To say that material relations of implication and incompatibility are not in general monotonic is to say that they are not *always* monotonic, not that they are *never* monotonic. The nonlogical implication of "Monochromatic swatch #13 is red," by "Monochromatic swatch #13 is crimson," and the nonlogical incompatibility of both with "Monochromatic swatch #13 is green," continue to hold when arbitrary further collateral premises are added. The same is true for many implications and incompatibilities involving claims about Euclidean plane figures, organisms classified according to Linnaean nomenclature, minerals specified by their chemical composition, and a host of other subject-matter specific reason relations.[9] This is what made it tempting and plausible for mid-twentieth century philosophers to talk about the "logic" of color-talk and other vocabularies that are not themselves logical vocabularies. Understanding ordinary empirical descriptive concepts (among others) involves not only having some practical ability (however partial and fallible) to distinguish good from bad implications and incompatibilities governing their possible applications, but also having some practical ability (however partial and fallible) to associate with them ranges of subjunctive robustness. That is the capacity to distinguish additional circumstances (premises) that would not defeat the implication or ruling out of the conclusion, from those that would. In navigating the complexly structured constellation of reasons for and reasons against in the course of defending and challenging various nonlogical claimables, the cases where the ranges of subjunctive robustness of the reason relations are *total* provide essential landmarks by which discursive practitioners can orient themselves. These are local regions of *persistent* implications and incompatibilities: ones that can be counted upon to remain good in the face of further collateral information or commitment.

The fact that these islands of monotonicity (or transitivity) of material reason relations are in general surrounded by seas of nonmonotonic implications and incompatibilities suggests another dimension of the expressive task of making explicit reason relations that expressivists about logic might reasonably call on logical vocabulary to perform. In addition to being able to use conditionals and negated sentences of a logically extended vocabulary to codify particular implications and incompatibilities that hold in the material base vocabulary to which logical vocabulary has been added (and from which its reason relations have been elaborated), it would be good to be able explicitly to mark regions of such reason relations where various elements of closure structure hold locally, even though they do not hold globally. In the next chapter, we will show how this can be done quite generally. For now it is enough to introduce the general idea.

Suppose we mark persistence of a particular implication by an upward arrow attached to the snake turnstile. If it is not only the case that $\Gamma \mid\!\sim A$, but also that for any and all sets of sentences Δ, $\Gamma \cup \Delta \mid\!\sim A$, then we can express that by writing $\Gamma \mid\!\sim^{\uparrow} A$. That local fact about the structure of implication in the base vocabulary can then be made explicit in a logically extended vocabulary if we add a new bit of logical vocabulary with a definition parallel to those used to illustrate the desired expressive role of the conditional and negation:

- Persistence-Codifying Modal Operator:
  $\Gamma \mid\!\sim \Box A$ if and only if $\Gamma \mid\!\sim^{\uparrow} A$.

If Γ does not just imply A, but persistently implies A, then we say that it also implies $\Box A$. That symbol is chosen because it can be read as "necessarily"—in the sense of "come what may," or "no matter what else is true." It explicitly marks the fact that the range of subjunctive robustness of the implication is *total*. The implication continues to hold no matter what additional premises are added to it. The box explicitly marks the monotonicity of the implication.

It is worth looking a little more closely at the sense in which the job of expressing explicitly the local structural feature of persistence or monotonicity of particular implications can be done by introducing logical vocabulary with a distinctively modal flavor. For it brings into view an important sense in which the corresponding principle DD we have suggested for thinking about the expressive role of conditionals defines them, too, as *intensional* connectives. If we fix the lexicon of a base vocabulary, we can think of the set of sets of sentences of that lexicon as a universe of points of semantic evaluation, playing a role analogous to that of possible worlds. Each such potential premise-set can be thought of as surrounded by a cloud of further sentences: those that are implicitly included in or excluded by that premise-set, in the sense of being implied by or incompatible with it: $\{X : \Gamma \mid\!\sim X\}$ and $\{Y : \Gamma \# Y\}$. The logical connective $\rightarrow$ is intensional in that the rule for introducing conditional sentences says that in order to determine whether or not $\Gamma \mid\!\sim A \rightarrow B$, one cannot just look at what is "true" or "false" at the point of evaluation Γ—in the sense of what is implied by or incompatible with Γ. Rather, one must look at what is "true" or "false," implied by or incompatible with, another, different, neighboring point of semantic evaluation, namely $\Gamma \cup \{A\}$. If B is "true at" or implied by that premise-set, then the logically complex sentence $A \to B$ is implied by Γ. In one core usage, the difference between extensional and intensional vocabulary is whether the status (paradigmatically, truth) at one point of semantic evaluation of sentences formed using that vocabulary is a function just of that same point of semantic evaluation, or whether to compute that semantic value at one point of evaluation, one must consult other points of semantic evaluation. The monotonicity-codifying box operator is (like a Ramsey-test conditional satisfying the DD condition) an intensional operator in this sense. For whether or not $\Gamma \mathrel{|\!\sim} \Box A$ depends on whether A is a consequence, not just of Γ, but also of all the supersets of Γ. This modal operator illustrates the principle that logical vocabulary can be called on to make explicit not just which particular reason relations a given premise-set stands in, but also facts about the structural conditions that hold locally, in its neighborhood, even if they do not hold globally.

Material incompatibility, we have argued, is symmetric—unlike implication. $\Gamma, A\#B$ if and only if $\Gamma, B\#A$. It follows that facts about the incompatibility of sentences with premise-sets are equivalent to facts about the incoherence of premise-sets. The two incompatibility facts just instanced are equivalent not only to each other, but to the incoherence of the set $\Gamma \cup \{A\} \cup \{B\}$. We can introduce a special symbol, $\bot$ (pronounced "perp") appearing only (alone) on the right of the turnstile, marking the set that appears on the right as incoherent: $\Gamma, A, B \mathrel{|\!\sim} \bot$. The same effect is achieved by using an empty right-hand side: $\Gamma, A, B \mathrel{|\!\sim}$. In either case, incompatibilities are traded for the incoherence of sets, and the incoherence of sets is expressed using the turnstile of implication. This common notational convenience has tended to lead—as a matter of psychology rather than philosophy—to a failure to acknowledge the co-equal status of incompatibility with implication as a basic form of reason relation. That is why we have, until now, instead used the nonstandard "#" notation to express incompatibility. In the fond hope that the temptation to misunderstand the significance of the notational convenience of expressing material incoherence of sets with the snake turnstile of nonmonotonic implication can be resisted, we will henceforth sometimes avail ourselves of that facility.

The fact that material incompatibility is not always monotonic, then, is another way of saying that material incoherence of sets of claimables is not monotonic. The fact that Γ is incoherent does not entail that $\Gamma \cup \Delta$ is incoherent. In the general case, rational incoherence (incompatibility) need not be persistent. If in some case it is not, if $\Gamma \mathrel{|\!\sim} \bot$ and it is not the case that $\Gamma, \Delta \mathrel{|\!\sim} \bot$, then the incoherence is curable, or defeasible. "Tweety is a bird" and "Tweety cannot fly" are incompatible, but adding "Tweety is a penguin" restores coherence. For a more serious example, wave behavior and particle behavior are incompatible in classical mechanics. In quantum mechanics, with lots of other auxiliary hypotheses added, they become compatible. Of course, *some* incoherences *are* persistent, in that not only is Γ incoherent, but so are all of its supersets. The combination of the claim that plane figure *T* is a Euclidean triangle and the claim that the angles of *T* add up to more than two right angles is incoherent (the claims are incompatible). And no further premises can cure that incoherence. Traditional logics, which treat both implication and incompatibility/incoherence as having full topological closure structure (so, as monotonic), enshrine a particularly striking relationship between the two reason relations. This is the principle "*ex contradictione quodlibet*" or explosion: from a contradiction, everything follows. As a structural principle regarding material reason relations of implication and incompatibility/incoherence generally, it is hard to know what to say in favor of this principle. It does not seem to correspond to anything in our ordinary reasoning practices, nor their institutionalized forms in scholarship, engineering, theoretical science, clinical medicine, or the law. This is a point on which structural logicism clearly seems to lead us astray. (That is why this principle is a permanent source of classroom embarrassment to teachers of elementary logic.)

However, the distinction between *persistently* incoherent premise-sets and those that are only curably or defeasibly incoherent is invisible to traditional logic, because of its commitment to the monotonicity of reason relations. Awareness of the distinction opens up the possibility of acknowledging that for ordinary, nonpersistently incoherent sets, we can still distinguish—as indeed, we do—between what follows from them and what does not. One might maintain that distinction even in cases where one does not have any idea what further premises might in principle restore coherence, were they added to the premises actually on hand. The fact that traditional logic offers no help in understanding this fact about them should not blind us to the possibility and indeed actuality of material base vocabularies in which at least potentially curably incoherent premise-sets do not explode implicationally. Expressivists about logic should accordingly aspire to construct logics in which incoherent premise-sets (except perhaps the incorrigibly, that is, persistently, incoherent—for instance, logically inconsistent—ones) are as tractable as coherent ones are. In Chapter Three we will see how this can be done—and further, how local regions of incoherence of premise-sets can explicitly be marked in logically extended base vocabularies where not all incoherence is incorrigible.

We can say a bit more about the relations among the different components of topological closure structure. It is widely recognized that failures of monotonicity (MO) generate failures of simple transitivity. So, a standard example of an MO failure is:

- Tweety is a bird, so Tweety can fly.

But *not*

- *Tweety is a bird and Tweety is a penguin, so Tweety can fly.

The corresponding failure of simple transitivity is

- Tweety is a penguin, so Tweety is a bird.
- Tweety is a bird, so Tweety can fly.

But *not*

- *Tweety is a penguin, so Tweety can fly.

Ryan Simonelli of our ROLE group has observed further that there is a general procedure for turning examples of failures of monotonicity into examples of failures of *cumulative* transitivity (CT): cases where explicitating a consequence adds further consequences. It is *not* the case that

- *Tweety is a bird, so Tweety is a non-penguin.

But we do have

- Tweety is a bird, so Tweety can fly.
- Tweety is a bird and Tweety can fly, so Tweety is a non-penguin.

Adding the consequence (fly) to the premise-set (bird) that does not imply non-penguin yields a premise-set that *does* imply non-penguin. In this way we can see that where one finds nonmonotonicity in material consequence relations, one will also find failures of CT.

There is a further argument against imposing transitivity as a global structural constraint. This is the observation that CT (in the context of CO), together with an implication-codifying (double) Ramsey conditional—that is, one that satisfies the Deduction-Detachment condition—forces MO (Hlobil, 2016). For if we start with some arbitrary implication $\Gamma \mid\!\sim A$, we can derive $\Gamma, B \mid\!\sim A$ for arbitrary B—that is, we can show that arbitrary additions to the premise-set, arbitrary weakenings of the implication, preserves those implications. And that is just monotonicity. For we can start with $\Gamma \mid\!\sim A$ and the CO-instance $\Gamma, A, B \mid\!\sim A$ and argue:

$$\dfrac{\dfrac{\Gamma \mid\!\sim A \qquad \dfrac{\Gamma, A, B \mid\!\sim A}{\Gamma, A \mid\!\sim B \to A}\text{[Ramsey Condition Right-to-Left]}}{\Gamma \mid\!\sim B \to A}\text{[CT]}}{\Gamma, B \mid\!\sim A}\text{[Ramsey Condition Left-to-Right]}$$

As a result, a proper conditional (one that satisfies the Deduction-Detachment condition) cannot be introduced conservatively on a nonmonotonic base. Such a conditional cannot in principle *explicate* (conservatively express) a nonmonotonic base. If we want such a conditional (and CO), we must forego CT as a global principle. For it is not just that CT creates a problem (forces monotonicity) if the language already contains an implication-codifying conditional. The problem is that we cannot *add* such a conditional to a nonmonotonic base language without endorsing new (monotonic) implications involving only the old, prelogical vocabulary. And that violates the explicative, implication-codifying expressive task characteristic of conditionals.

The idea that reason relations in general do, or even must, share the topological closure structure of specifically *logical* reason relations is a restricted kind of logicism: *structural* logicism about reasons. The nonmonotonicity (and, as we will see, also nontransitivity) of nonlogical implication relations—the fact that they are not necessarily or in general monotonic or transitive—shows that the structural logicist thesis is false. That fact poses a challenge to the more general logicist program of reconstructing all good reasoning on the basis of logically good reasoning: logically valid implications and logical inconsistencies. If structural logicism about material reason relations is not true, then neither is logicism about reasons generally—at least not in its original form. That challenge need not be understood as fatal to the general logicist program. Perhaps nonmonotonic implications can somehow be reconstructed out of monotonic ones. Something like that is what traditional nonmonotonic logics aspire to: to start with proper, monotonic logical systems, and tweak them somehow (add structure such as preference orderings, distinguish different classes of premises, restrict models) so as to yield nonmonotonic reason relations.

But this structural difference between logical reason relations and reason relations more generally presents a problem for expressivism about logic just as much as it does for logicism about reasons. For it is not at all obvious how monotonic logical vocabulary could make explicit nonmonotonic implications and defeasible incompatibilities. Expressivists have a project here that is different from that of traditional nonmonotonic logicians. The important task is not seen to be constructing nonmonotonic *logics*, but rather logically codifying nonmonotonic (and nontransitive) consequence relations (and incompatibilities). It turns out, as we'll see in the next chapter, that to do that one does *not* need logics that are themselves nonmonotonic (or nontransitive). Though it sounds paradoxical, logics that are expressively adequate to make explicit radically substructural, open material reason relations can themselves exhibit the full topological closure structure characteristic of classical logic.

To see how this can be possible, one must distinguish two different notions of consequence (and incompatibility) that govern the use of logical vocabulary. First are the implication relations that hold among sentences of some particular logically extended base vocabulary. If the reason relations of the base vocabulary are themselves structurally open (nonmonotonic or nontransitive), as we have argued they generally are, then since the reason relations of the logically extended vocabulary elaborated from them must be conservative, those reason relations, too, must be structurally open. If $p \mathrel{|\!\sim} q$ holds in the base vocabulary and $p, r \mathrel{|\!\sim} q$ does not hold in the base vocabulary, then $\mathrel{|\!\sim} p \rightarrow q$ but not $\mathrel{|\!\sim} (p \wedge r) \rightarrow q$ will hold in the logical extension of *that* base vocabulary. Note that in nonmonotonic settings, the fact that $p \rightarrow q$ is implied by the empty set of premises, and is in that traditional sense a theorem, does *not* entail that it is implied by *every* set of premises. For instance in the situation being envisaged, the fact that $\{p, r\}$ does not imply q does not entail that r does not imply $p \rightarrow q$. The notion of "theorem" that matters in nonmonotonic settings is being implied by *every* premise-set, rather than being implied by the empty premise-set. Global structural monotonicity runs these two distinct notions together.

But that failure of monotonicity in the logical extension of some base vocabularies is entirely compatible with the *purely logical* implications, the implications that hold in virtue of logical form alone, being fully monotonic. For the sequent codifying a nonmonotonic implication is derivable only in proof trees that have as leaves sequents of the particular base vocabulary in which p materially implies q and $\{p, r\}$ does not materially imply q. Purely logical implications, implications that hold in virtue of logic alone—that is, the connective rules (plus Containment)—hold regardless of what base vocabulary one uses those connective definitions to extend, including those in which p does not materially imply q and $\{p, r\}$ does materially imply q. The expressive logic we will recommend, NMMS (and, indeed, the even more expressively capable non-contractive version of NMMS), is capable of making explicit arbitrary, radically substructural base vocabularies: those that are in general neither monotonic nor transitive. But it is itself not a nonmonotonic or nontransitive logic. It is not only supraclassical, there is a sense in which it just *is* (a formulation of) classical logic.

2.5 Even More Radically Open Structures of Reason Relations

So relations of consequence (following from or being a reason for) in general do not have the full structure of closure operations—as the tradition of Tarski and Gentzen takes it that specifically *logical* consequence relations not only do, but must. In the same way, and for the same reasons, material relations of incompatibility are not in general monotonic. In view of these features of ordinary, nonlogical reasoning, expressivists about logic should aspire to codify open-structured reason relations, including nonmonotonic and nontransitive ones. That is to give up *structural* logicism, and to remove much of the ground supporting logicism about reasons generally. One question that then arises is whether there are any structural conditions weaker than the full topological closure structure that *can* be counted on to characterize material reason relations.

Containment (CO) says that in the case of consequences of implications that are also premises of those implications, one can weaken the premise-set with arbitrary additional premises without infirming the conclusion. It is plausible that among the consequences of any premise-set are those premises themselves. Relevance logicians, however, object. They do not recognize as good, implications that involve premises that are irrelevant to the conclusions. That is a reason to treat the addition of irrelevant premises as able to defeat otherwise even trivially good implications such as instances of Reflexivity, of the form $A \mid\!\sim A$. Whether or not the relevantists are right about that, logical expressivists ought to aim for a logic that does not build in the presupposition that they are wrong. Put another way, a logic that could codify even consequence relations that fail to satisfy Containment is clearly more expressively powerful (other things being equal) than, and so preferable to, one that cannot. The logic NMMS introduced in Chapter Three works just fine (is provably LX) even for base vocabularies whose consequence relations do not satisfy Containment—though, as we will see, the result is that the logical consequence relations defined are no longer supraclassical.

Denying Monotonicity (MO) means not assuming that one can weaken *every* implication with arbitrary additional premises without infirming the conclusion. Is there a way to specify some more restricted set of additional premises one could add that *would* be guaranteed not to defeat an implication? One plausible candidate answer to this question is given by what has been called "*cautious* monotonicity." Cautious monotonicity says that while one might indeed not be able to weaken an implication by adding just any sentence as a collateral premise without running the risk of infirming it, it should at least be safe to add further premises that are already implied by the original premise-set.

Cautious Monotonicity (CM):

$$\frac{\Gamma \mid\!\sim A \qquad \Gamma \mid\!\sim B}{\Gamma, A \mid\!\sim B}$$

The idea is that since the premise-set Γ already implies A, adding A as an explicit premise should not cause any trouble with other consequences of Γ. Even though there might be *some* additional premises that *would* infirm the implication, sentences that are *already implied* by the premise-set are not among them.

It has often been argued not only that cautious monotonicity is a plausible principle, but that it is in effect indispensable: that it is a *minimal* condition that well-behaved nonmonotonic consequence relations must satisfy.[10] Satisfying CM is generally regarded as a criterion of adequacy for assessing nonmonotonic logics. CM plays a prominent role, for instance in what Kraus, Lehman, and Magidor call the "core properties" or the "conservative core" of non-monotonic systems (and for this reason are now often called the "KLM structural properties" required of nonomonotonic systems), and count it a signal virtue of their preferential semantics for nonmonotonic logic that it validates this structural principle (Kraus et al., 1990). By contrast to monotonicity, there has not been much skeptical philosophical attention directed at cautious monotonicity. This is a shame, because the underlying issues here are just as important, and addressing them is deeply revealing of considerations that remain invisible if the discussion remains at the level of the much stronger structural principle of monotonicity.

The first step in appreciating this is realizing that cautious monotonicity is the dual of cumulative transitivity, a version (the shared context version) of Gentzen's "Cut."[11] This structural principle is expressed in Tarski's algebraic metalanguage for consequence relations by the requirement that the consequences of the consequences of a premise-set are just the consequences of that premise-set, and by Gentzen as the principle that adding to the explicit premises of a premise-set something that is already part of its implicit content does not add to what is implied by that premise-set. It is the principle appealed to in chaining together implications in extended consecutive reasoning.

Cumulative Transitivity (CT):

$$\frac{\Gamma \mid\!\sim A \qquad \Gamma, A \mid\!\sim B}{\Gamma \mid\!\sim B}$$

CT says that adding a consequence of a premise-set to that premise-set never *adds* consequences—that what a premise-set implies when we add to it its own consequences, it already implies all on its own—while CM says that adding consequences to the premise-set never *subtracts* consequences the original premise-set had.

Here is a way to think about the underlying issue. Using language that was second nature to Leibniz and Kant, we can think about the *content* of a set of claimables in the literal sense of what is *contained in* it. A premise-set $\Gamma = \{A_1, A_2, \ldots, A_n\}$ literally contains all of the sentences A_i in the set-theoretic sense that these are the elements of the set Γ. We may say that it contains them *explicitly*, since they are what we specify when we specify the set. As so contained, they are the *explicit content* of the set—"content" etymologically being what is contained. If it now happens that Γ implies *A*—in our notation, $\Gamma \mid\!\sim A$—then we can say that *A* is *implicit* in Γ, in the literal sense of being *implied by* it. *A*, then, is part of the *implicit content* of the premise-set Γ. (Analogously, we might think of every set Δ that is materially *incompatible* with Γ as being part of Γ's implicit *contrastive* content.)

Then CM and CT can be thought of as having a common topic. Both concern what happens when the status of some consequence of a premise-set is changed, by turning it into an additional premise. The process of moving a sentence from the right-hand side of the implication turnstile to the left-hand side, from appearing as a conclusion to appearing as a premise, might be called the process of *explicitation*. For it is the process of making some *implicit* bit of content *explicit*, turning what is *implicitly* contained in a premise-set (that is, what is implied by it) into a premise that is explicitly contained in the (new) premise-set.[12] Explicitation in this sense is not at all a *psychological* matter. And it is not even yet a strictly *logical* notion. For even *before* logical vocabulary has been introduced, we can make sense of explicitation in terms of the structure of non-logical, *material* consequence relations. Noting the effects on implicit content of adding as an explicit premise sentences that were already implied is already a process available for investigation at the methodological level of the *prelogic*, in the base vocabulary, before that material base is logically extended.

Both CT and CM concern the effects that explicitation has on the *consequences* of the premise-set, comparing the consequences before explicitation with the consequences after explicitation. Since the consequences of a premise-set are its implicit content, CT says that explicitation does not *gain* any implicit content, and CM says that explicitation does not *lose* any implicit content. CT says no consequences are added, and CM says no consequences are subtracted. Together, they entail that ***explicitation is inconsequential***: making explicit some or all of the implicit content of a premise-set has no effect on its consequences at all. Moving a sentence from the right-hand side of the implication-turnstile to the left-hand side does not change the consequences of the premise-set. It makes no difference whatever to the implicit content, to what is implied.

We began by opposing an *expressivist* approach to understanding the relations between logic and reasoning (mediated by reason relations of implication and incompatibility) to a *logicist* approach to understanding them, according to which all good reasons are *logically* good reasons: every genuine implication is valid in virtue of the logical form of its premises and conclusion. We then considered a weaker, purely *structural* form of logicism. It claims that the algebraic structure of material reason relations of implication and incompatibility is the same as the algebraic structure of specifically *logical* relations of implication and incompatibility. For historical reasons we have gestured at, philosophers of logic have taken that algebraic structure to be topological closure. Topological closure is a matter of satisfying monotonicity and transitivity, MO and CT (as well as Containment, CO, or at least Reflexivity, RE). The present claim is that it is more philosophically revealing to focus on a different kind of closure structure, which involves pairing CT not with MO, but with CM. This might be called "*explicitation closure*," since it entails that explicitation is inconsequential. Since MO entails CM, rejecting the explicitation-closure form of structural logicism—by denying that explicitation is inconsequential for material consequence relations—will entail rejecting the topological-closure form of structural logicism. Consequence relations for which CT fails are *nontransitive*. Implication and incompatibility relations for which MO fails are *nonmonotonic*. Reason relations for which CM fails—implications and incompatibilities can be defeated even by the addition of premises that are already implied by a premise-set—are even more radically substructural or structurally open. They are *hypernonmonotonic*. We should understand that material reason relations in general can include hypernonmonotonic implications.

It might well be sensible to require the inconsequentiality of explicitation as a structural constraint on *logical* consequence relations. But just as for the logical expressivist there is no good reason to restrict the rational relations of implication and incompatibility we seek to express with logical vocabulary to monotonic or transitive ones, there is no good reason to restrict our expressive ambitions to consequence relations for which explicitation is inconsequential. On the contrary, there is every reason to want to use the expressive tools of logical vocabulary to investigate cases where explicitation *does* make a difference to what is implied.

One such case of general interest is where the explicit contents of a premise-set are the records in a *database*, whose implicit contents consist of whatever consequences can be extracted from those records by applying an *inference engine* to them. (The fact that the "sentences" in the database whose material consequences are extracted by the inference engine are construed to begin with as *logically* atomic does not preclude the records having the "internal" structure of the arbitrarily complex datatypes manipulated by any object-oriented programming language.) It is by no means obvious that one is obliged to treat the results of applying the inference-engine as having exactly the same epistemic status as actual entries in the database. A related case is where the elements of the premise-sets consist of experimental *data*, perhaps measurements, or observations, whose implicit content consists of the consequences that can be extracted from them by applying a *theory*. In such a case, explicitation is far from inconsequential. On the contrary, when the CERN supercollider produces observational measurements that confirm what hitherto had been purely theoretical predictions extracted from previous data, the transformation of rational status from *mere* prediction *implicit* in prior data to actual explicit empirical observation is an event of the first significance—no less important than the observation of something incompatible with the predictions extracted by theory from prior data. This is the very nature of empirical *confirmation* of theories.

Imposing Cut and Cautious Monotonicity as global structural constraints on material consequence relations amounts to equating the epistemic status of premises and conclusions of good implications. But in many cases, we want to acknowledge a distinction, assigning a lesser status to the products of risky, defeasible inference. In an ideal case, perhaps this distinction shrinks to nothing. But we also want to be able to reason in situations where it is important to keep track of the difference in status between what we take ourselves to know and the shakier products of our theoretical reasoning from those premises. This idea is one important motivation for Strict-Tolerant logics, paradigmatically, ST (Cobreros et al., 2013). Chapter Five discusses how this logic shows up from the vantage point of the more expressively powerful treatment of conceptual roles made possible by our implication-space semantic metavocabulary. Our expressive aspiration is for a logic that can codify reason relations that do not presuppose or build in a structural requirement of explicitation closure: the demand that explicitation always be inconsequential.

2.6 Explicitation Closure and Rational Hysteresis

Let us take stock. We began by considering two diametrically opposed approaches to understanding the relations between logic and reasoning. Taking on board the idea that logic concerns the reason relations of implication and incompatibility that govern reasoning practices and processes sharpened the issue somewhat. Logicism claims that logic determines the proper relations of implication and incompatibility: that implications hold rationally just in case they are or are supported by *logically* good implications (valid deductions), and that rational incompatibilities are or are supported by *logical* incompatibilities, that is, formal inconsistencies. Expressivism understands the task of logic to be expressing material, prelogical reason relations of implication and incompatibility: making them explicit in the sense of sayable, claimable contents, for which reasons can be asked and given. The rational importance of logical vocabulary is not that it lets us prove theorems of logic that determine what relations of implication and incompatibility really hold, but that the sentences of any logically enriched prelogical base vocabulary let us say what relations of implication and incompatibility hold in that prelogical base vocabulary.

We then considered a weaker form of logicism: *structural* logicism. This is the view that the algebraic structure of material reason relations is the same as the algebraic structure of specifically logical reason relations: implications and incompatibilities that hold just in virtue of the logical form of the sentences involved. For the case of logical consequence, it is generally agreed that it has the structure of a topological closure operation. Combining Tarski's and Gentzen's versions of these structural principles (so as to extrude irrelevant details particular to their formulations), this means that consequence relations satisfy Containment (CO), Monotonicity (MO), and Cumulative Transitivity (CT). The traditional form of structural logicism accordingly takes implication to have a *topological closure* structure. Assuming Tarski and Gentzen are right about the structure of logical consequence (and this much is not at issue, for instance, between classicists and intuitionists), logicism about implication generally entails the topological closure form of structural logicism. If structural logicism is wrong about implication in general, if material consequence relations do not in general exhibit the topological closure that comprises CO, MO, and CT, then logicism about reasons in general cannot be right either.

We then argued that, however it might be with specifically *logical* implication, *material* consequence relations are not in general monotonic. That is enough to show that the topological closure version of structural logicism is not true. But what structure do consequence relations in general exhibit? It was pointed out that a popular principled fallback from MO is Cautious Monotonicity (CM).[13] It is the principle that no consequences of a premise-set Γ are lost by adding any collateral premises that are already consequences of Γ. CM seems a particularly natural weakening of MO to consider, because it is dual to CT. CT says that adding consequences of Γ to Γ never *adds* any new consequences, and CM says that adding consequences of Γ to Γ never *subtracts* any consequences. Observing this duality brings into view the operation they share: moving a sentence across the implication turnstile, from being a conclusion to being a premise. Thinking of what a premise-set implies as what it contains *implicitly* (its implicit content) and the actual elements of the premise-set as what it contains *explicitly* (its explicit content) makes it natural to call this process "explicitation." It is making implicit content explicit: a prelogical sort of expression.

Explicitation, in turn, makes visible a further kind of closure structure: explicitation closure. If both CM and CT hold globally for an implication relation, then explicitation is guaranteed to be *inconsequential*. Making implicit content explicit never affects implicit content, neither increasing or decreasing it. This sort of closure is weaker than topological closure, just insofar as CM is weaker than MO. But it, too, expresses commitment to a kind of *stability* of consequences, a kind of closure.

And with this weaker sort of closure structure comes a new sort of structural logicism: *explicitation-closure* structural logicism (now distinguished from *topological*-closure structural logicism). This is the claim that consequence relations in general have at least this much of the structure of logical consequence: CM and CT, as well as CO.[14] Explicitation closure might indeed seem to be an attractive fallback from topological closure as a candidate for being the structure of consequence relations generally. It is implied by the full topological closure Tarski and Gentzen take to be characteristic of logical consequence, and would permit acknowledgment of a structural distinction between logical and material consequence relations, if the material ones satisfy only the weaker explicitation-closure structure, rather than the full topological-closure structure of logical consequence as traditionally conceived.

To argue against the explicitation-closure form of structural logicism we introduced a special case of consequential reason relations to serve as a model. It can be helpful to think of the explicit premises of an implication as a database, and the turnstile as standing for a theory functioning as an inference-engine that, when applied to the database, yields the conclusions, thereby extracting the content implicit in the database. Thinking of instances where the database contains observational data and the inference-engine extracts the predictions of some theory shows that the inconsequentiality of explicitation is *not* a condition we want to insist on for implication or consequence relations in general. The difference in status between what has been observed or measured and what is merely predicted by theory is too important to have the boundary between them structurally erased. Explicitation is not in general or *de jure* inconsequential: it can make a difference to what is implied by a premise-set. The explicitation-closure conditions CM and CT together imply that explicitation is inconsequential in this sense. So this weaker form of closure should also be rejected as a global structural constraint on consequence relations, and explicitation-closure structural logicism must accordingly be rejected along with topological-closure structural logicism.

Remember the issue that led to addressing the issue of whether material consequence relations should be understood to be structurally closed in the first place (whether in the topological or the explicitation sense). It was to understand the constraints on *expressing* material reason relations using *logical* vocabulary. The database + inference engine model reminds us that we should aim for the greatest possible flexibility and expressive power possible. We should build in as few *a priori* constraints on logically codifiable inference engines as possible. The expressivist's aim should be to produce and deploy logical tools for expressing reason relations of all intelligible structures. The expressivist ideal is to develop the expressive power to make explicit *any* and *all* species of the turnstile, any and all senses of "follows from." Thinking in terms of databases and inference engines reminds us of just how capacious that class (and so that aspiration) is. Perhaps it is a utopian aspiration. (Spoiler: it is not.) From this point of view, CM should not be assumed to hold globally, any more than CT should. From an expressivist point of view, we want a logic that can be introduced conservatively over, and has the expressive power to codify, material reason relations that are *open*, not just in not being *topologically* closed, but in not being *explicitation* closed, either. Structural logicism in both forms must be rejected.

The material consequence relations that the expressivist takes it logic should aim to codify are accordingly radically substructural, in that it should not be presupposed that material, nonlogical relations of implication satisfy global structural principles of the sort characteristic of specifically *logical* consequence relations. The kind of structure denied is *closure* structure, of the two sorts distinguished here: topological closure and explicitation closure. The substructural consequence relations they contrast with are *open*, in both corresponding contrasting senses. The significance of reason relations with open structure is best grasped by thinking about the process of explicitation. Explicitation as a process is *drawing* or extracting consequences, and adding them to the premises from which one reasons. This is a central, indeed, essential kind of *inference*: acknowledging *explicitly* (by treating them as premises in further implications) conclusions that were *implicit* in some prior premise-set—"implicit in" in the literal sense of "implied by." (Harman reminds us that this is not the only rational inferential process governed by implication relations. Denying what is in this sense implicit in a premise-set can also entitle one explicitly to deny some or all of the premises.) The key point is that in an open structural setting, making explicit any set of consequences of a premise-set might add some new consequences (where there are local failures of CT) and subtract others (where there are local failures of CM), relative to the consequences of the original premise-set. This has a number of consequences, which highlight the striking differences between the general case and structurally closed, specifically *logical* consequence relations.

Most importantly, realizing that explicitation is not in general inconsequential, that failures both of CM and of CT are to be expected, challenges the common philosophical conception of the *rational* closure of a knower's beliefs. It has seemed natural to some to think that rationality requires that one believe all the consequences of one's beliefs or commitments: everything that follows from them. In a familiar dialectic, opponents point out that this is an impractical demand. There are far too many such consequences, and simply surveying them would take too long. In response, while it is acknowledged that it might be impossible actually to achieve the rational closure of one's beliefs or commitments (attitudes of accepting or denying claimables), so that treating doing so as constitutive of rationality would entail that no-one ever is or could count as fully rational, nonetheless it is suggested that such closure could still be understood as a regulative ideal of rationality. It is then objected (for instance, by Harman) that we should not understand reasoning, in the sense of drawing further consequences from our beliefs, as governed by a norm we can never in practice live up to. The important thing to realize about this debate is that it assumes that the concept of all the consequences of a premise-set (its rational closure in the sense of everything that follows from it) makes sense or has a definite reference. The argument has been about whether closing one's beliefs under rational consequence is practical, desirable, a regulative ideal of rationality, and so on. But if CM or CT sometimes fail, then there is in general no such thing as *the* set of consequences of a premise-set. The dialectic only makes sense assuming at least the weaker form of structural logicism: that, as for purely logical implication, explicitation is in general inconsequential.

If material reason relations can be hypernonmonotonic and nontransitive, then rational consequence does not have a closure structure. There need be no closed set of consequences for each premise-set, which cannot be added to or subtracted from by continuing to extract consequences. And that means that the process of explicitation—of explicitly acknowledging consequences of the claimables one accepts, and using them as premises in the drawing of further consequences for which they provide reasons—becomes an open-ended enterprise. There is no unique endpoint that one is guaranteed eventually to reach, by just keeping at the process of explicitating rational consequences indefatigably, and making no mistakes.

- Acknowledging some consequences one's commitments give one reasons for can put one in a position where one no longer has sufficient reasons to draw other consequences that one could otherwise have drawn (that one would otherwise have been entitled to draw) from the original set.

That is a consequence of the failure of CM to hold globally.

- And explicitating to begin with only some of the consequences of the original premise-set might well provide reasons for new consequences, that the original premises, taken all together, did not imply or directly provide reasons for.

That is a consequence of the failure of CT to hold globally.

If explicitation closure is not guaranteed by the structure of implication as such, then there is no guarantee that the process of explicitating consequences, and then explicitating the consequences of *that* expanded premise-set, and then explicitating the consequences of *that* set, and so on, will reach a fixed conclusion. In open settings one cannot be sure in advance that the process of explicitation will reach a stable stopping-place—that it will arrive at a premise-set all of whose explicitations are inconsequential, involving no violations of CT or CM. It can happen that every position $X \cup Y$ one arrives at by explicitating consequences of the original premise-set X (CO guarantees one will arrive only at supersets of X) still has some implicit content, some set of consequences, such that when they are added to $X \cup Y$ as explicit premises, the result is a different set of consequences than $X \cup Y$ had. The process of explicitation need not end. The explicitation closure conditions, by contrast (like the topological closure conditions), guarantee *finality*: that $Cn(X) = Cn(Cn(X))$ for every premise-set.

A further observation is that with open consequence relations there is nothing privileged about the result of explicitating *all* of X's consequences at once. One might get a larger implicit content by explicitating only some of the consequences of a premise-set. Closed consequence relations ensure that the result of explicitating all the consequences of a premise-set will include the result of explicitaing only some of them. So not only is finality guaranteed, but one can get to the final fixed point of the explicitation process in a single step. Explicitation of topologically closed consequence relations is *immediate*.

Since finality is not guaranteed, explicitation of consequence relations with open structure can be radically path-dependent. Nor is there any guarantee that following any particular explicitation path that starts by adding those consequences will ever arrive at any of the premise-sets or consequence sets reached by starting with the explicitation of a different proper subset of the consequences of Γ. As a result, the consequences that come into view at any point in the process of explicitating some premise-set Γ depend on which of Γ's consequences one chooses to explicitate first. This is explicative *hysteresis*. By contrast, explicitation of structurally closed consequence relations is *stable*, in the sense of being path-independent. Since it is *final* there is a fixed endpoint to the process: $Cn(X)$. Since it is *immediate* one can jump to that endpoint in a single explicitating step. And since it is *stable*, if one *did* explicitate step-wise, the results would be the same endpoint, for all explicitation paths from X to $Cn(X)$.

Finality, immediacy, and stability of explicitation are all very useful features of consequence relations. There are many good reasons to want to build them into the *logical* consequence relations that govern the use of the logical vocabulary whose expressive task it is to make explicit the reason relations that govern the use of all kinds of vocabulary. (The purely logical consequence relation determined by the rules of our favored logic NMMS does have all these nice closure properties—though the full consequence relation of the logical extensions of arbitrary base vocabularies does not.) But if, through a thoughtless and misplaced commitment to either variety of structural logicism, we project those ideals onto the actual material consequence relations that govern the use of nonlogical vocabulary and articulate the conceptual contents they express, we make invisible essential structural features of the crucial rational *process* of *drawing* consequences from premise-sets, of acknowledging explicitly the implicit, consequential contents of explicit commitments. The process of explicitation is an important species of inference. Studying it is one of the reasons we want logical tools to make reason relations explicit. The most elementary sort of pragmatism counsels that we do not obscure rational processes and practices by imposing Procrustean *a priori* structural restrictions derived from mistakenly taking specifically *logical* reason relations of implication and incompatibility as the model on which to understand the structure of reason relations in general.

The fundamental structural feature of the process of explicitation that is being highlighted here is that sequentially extracting consequences from a premise-set—explicitly acknowledging them, and then extracting more of the implicit content of that set—is not a process that is guaranteed either to converge or, in the cases where it *does* converge, to reach a *unique* final set of consequences.

Reasoning in this explicative sense of making explicit what is rationally implicit in a set of commitments is in principle severely *path-dependent*. Where one ends up depends upon what choices one makes about what to explicitate *first*, and in general on the *sequence* of explicitations one effects. Indeed, those choices determine even whether an inferential process of extracting and acknowledging consequences converges at all, as well as where it ends up. The historical path-dependence of consequences and incompatibilities is a deep *structural* feature of material, nonlogical reasons as such.

In the physical sciences, the technical term for path-dependence of processes is "hysteresis." That is why the phenomenon in question here is called the *rational hysteresis* of explicitation—the hysteresis of making explicit or acknowledging explicitly what is rationally implicit in (implied by) a set of commitments. It ensures that *history* is an intrinsic feature of *reason*. This is the *rational* (not yet logical) basis of the *historicity of reason*. It is an essential element of the explanation of why *discursive* creatures, creatures who can *talk*—that is can engage in practices of making claims and challenging and defending them—have *histories*, and not just *natures*.

It was pointed out above that material incompatibility relations can be nonmonotonic, just as material consequence relations can be. Material incompatibility relations can also be hypernonmonotonic. Base vocabularies can have premise-sets for which not only MO fails, but Cautious Monotonicity fails. These are sets of sentences whose status as coherent or incoherent can be changed by explicitation of some or all of their consequences. Perhaps the most interesting case is coherent sets that are implicitly incoherent, in that adding one or more of their consequences as explicit premises results in an explicitly incoherent set. Many, perhaps most, philosophy books are probably like this. Their explicit content is expressed by the sentences that occur on their pages. Their implicit content is what those sentences jointly imply but do not explicitly state. Part of the work of readers, reviewers, and seminars studying the text is to extract at least some of its consequences and critically confront the explicit text with challenges presented by incompatibilities between the explicit content and what it implies. Charitable interpreters will consider whether the incoherences that emerge from such exercises are curable by the addition of further premises.

For the implicit incoherence of a premise-set need not be *persistent*. It might be curable or defeasible by the addition of further premises. A standard strategy for curing an incompatibility by adding a further premise is epitomized in the Scholastic maxim "When faced with a contradiction, make a distinction." (That is roughly how functionalists in the philosophy of mind proposed to reconcile the materialist's claim that thoughts must be specifiable in the language of physics or neurophysiology with the dualist's claim that what makes something the thought that *p* cannot be specified in the language of physics or neurophysiology.) The possibility under consideration is that the distinction required to reconcile incompatible claims might be implicit in those claims themselves. (The claim that, at least within some domains, *every* contradiction implicitly contains a distinction that would resolve it is a recognizably Hegelian one.) Some additional premises that would, if explicitly added to the premise-set, result in a coherent consequence set might sometimes be found even among those implied by the incoherent result of prior explicitation. Exploring the many paths in and out of the dangerous territory of explicit incoherence that explicitation of implicit content can trace out is an interpretive adventure. It is characterized by what we are in a position to recognize as hermeneutic hysteresis (not to be confused with hermeneutic hysteria). Among the way-stations along many such explicitation paths will be found explicitly incoherent sets of claims that are implicitly coherent. In such cases, the incoherence is corrigible by adding premises that are already implied by an explicitly incoherent premise-set. The difficulties of the text as it stands can be rectified by merely explicitly drawing consequences to which the text already leads. The best readers, editors, and referees will be most helpful when pointing these out.

Exploring the material reason relations that articulate the conceptual contents expressed by the use of nonlogical vocabularies by traversing explicative pathways leading from one premise-set to another is a principal form of semantic understanding—understood now as a process rather than a state. Expressivists want a logic expressively powerful enough to make explicit what we are doing and what we find out when engaged in this sort of hermeneutic enterprise. That requires distinguishing at least the three kinds of incoherence/incompatibility just pointed out: explicit, implicit, and persistent. The very existence of these important phenomena is rendered invisible by structural logicism. Traditional logics are inadequate to the task of expressing these hypernonmonotonic, open-structured regions of material reason relations.[15]

2.7 Conclusion: Expressivist Criteria of Adequacy for Logical Vocabulary

This chapter began by contrasting two different ways of thinking about the relations between logic and reasons in general. Logicism about reasons understands "good reason" to mean "logically good reason." Reasons for commitments are governed by logically valid implications, and reasons against commitments are governed by logical inconsistencies. Where it is not obvious on the surface that this is so for some nonlogical reasons, if reasons are to be understood to be involved in the use of that nonlogical vocabulary, a logical deep structure must be discerned as underlying its use. By contrast, expressivism about logic treats the material goodness of prelogical reason relations of implication and incompatibility as prior in the order of explanation to reason relations of specifically *logical* consequence and inconsistency. The defining task of logical vocabulary is an expressive one: to make explicit, in sentential, claimable form, the reason relations of material consequence and incompatibility that normatively govern the use of sentences expressing claimables in nonlogical base vocabularies. Logical reason relations are *elaborated from* the reason relations of base vocabularies, and the logically complex sentences whose use those logical reason relations govern are *explicative of* the implications and incompatibilities of those base vocabularies. We say that logical vocabulary is "LX" for the base vocabularies from which it is *el*aborated ("L") and of which it is *ex*plicative ("X"). An expressively ideal logical vocabulary would be *universally* LX: LX for *any* base vocabulary (lexicon of sentences governed by reason relations) whatsoever.

The bulk of this chapter has been concerned with a structural stumbling block that confronts both the logicist and the expressivist programs.

The algebraic structure of specifically logical reason relations is quite different from the algebraic structure of reason relations of implication and incompatibility in general. Logical consequence has a topological closure structure. It is monotonic, transitive, and reflexive (or satisfies Containment, sometimes called "Contexted Reflexivity"). Nonlogical implications, by contrast, can be nonmonotonic and nontransitive. Indeed, they can be hypernonmonotonic, in the sense that the implication of a conclusion by a premise-set can be defeated or infirmed by making some of its implicit content explicit: adding as an explicit premise something that was already implied by the original premise-set. Material incompatibilities can also be nonmonotonic, and even hypernonmonotonic. The logicist about reasons is obliged somehow to construct substructural (open-structured) reason relations out of the structurally closed relations of logical consequence and inconsistency. That is the task of traditional nonmonotonic logics.

This structural mismatch between material and logical reason relations also poses challenges for both sides of the expressivist ideal for logical vocabulary: that it should be capable of being elaborated from and explicative of the reason relations of every material base vocabulary. On the side of elaboration, if logically complex sentences using conditionals and negations are to be capable of expressing explicitly material relations of implication and incompatibility, their addition must conservatively extend the reason relations of the base vocabulary. If adding the new logical vocabulary changed the material reason relations of the base vocabulary, for instance by imposing further structure on them, that would undermine the capacity of the new logical vocabulary to express them explicitly. But some standard connective rules force the consequence relation governing the whole logically extended vocabulary to be monotonic, in the context of other plausible principles. For example, part of Gentzen's rule for introducing conjunction on the left side of sequents (put in the notation used here) is:

$$\frac{\Gamma, A \mathrel{|\!\sim} \Delta}{\Gamma, A \wedge B \mathrel{|\!\sim} \Delta}$$

If $\Gamma, A \wedge B \mathrel{|\!\sim} \Delta$ entails $\Gamma, A, B \mathrel{|\!\sim} \Delta$ (as seems plausible), then this permits a transition from $\Gamma, A \mathrel{|\!\sim} \Delta$ to $\Gamma, A, B \mathrel{|\!\sim} \Delta$ via the intermediate step of $\Gamma, A \wedge B \mathrel{|\!\sim} \Delta$, even if all of Γ, Δ, A, and B are sentences of the nonlogical base vocabulary. This just is monotonicity. In this case, one could avoid building in monotonicity into the connective definition by using a "multiplicative" rule rather than the "additive" one:

$$\frac{\Gamma, A, B \mathrel{|\!\sim} \Delta}{\Gamma, A \wedge B \mathrel{|\!\sim} \Delta}$$

If one imposes monotonicity as a global structural constraint, the additive and the multiplicative rules are equivalent. If not, they come apart—and the difference becomes important. The point remains: expressivist logics must be able to accommodate base vocabularies that have radically open structures: nonmonotonic, hypernonmonotonic, and nontransitive, without imposing additional structure on the original reason relations. The structural disparity of logical and material reason relations makes the conservativeness condition on the elaboration of logical reason relations from material base reason relations challenging to satisfy.

On the side of explication, expressivist logics aim to be capable of explicitly expressing (putting in statable, claimable form) the radically open-structured reason relations of any and all base vocabularies, even where they are hypernonmonotonic, nontransitive, and fail to satisfy other minimal global structural conditions such as Containment. That principle says that all sequents of the form $\Gamma, A \mathrel{|\!\sim} A, \Delta$ are good: that what a premise-set contains explicitly, it also implies, and so contains implicitly. That is a plausible, because severely restricted, monotonicity principle, allowing the weakening of instances of Reflexivity of the form $A \mathrel{|\!\sim} A$ by arbitrary Γ and Δ. But it is rejected by relevance logicians (due precisely to the irrelevance of Γ and Δ to the goodness of the implication), and expressivists aspire to be able to codify the reason relations of relevance logics, just like any other base vocabulary. Sufficiently expressively powerful logics should even be able to handle failures of Gentzen's principle of Contraction:

$$\frac{\Gamma, A, A \mathrel{|\!\sim} \Delta}{\Gamma, A \mathrel{|\!\sim} \Delta}$$

That principle is not satisfied by the consequence relation of Linear Logic. (Its inventor, Jean-Yves Girard says that anyone who accepts Contraction deserves "two kicks in the ass—not two occurrences of the same kick" (Girard, 2001, 482).) Expressivists want logics that can make explicit the reason relations of Linear Logic, just as for any other base vocabulary.

This is a tall order. Chapters Three and Five show how this expressivist aspiration can be fulfilled, nonetheless. Astonishingly, familiar logics can be tweaked so as to be demonstrably LX for all these sorts of radically substructural reason relations—and are supraclassical (for the multisuccedent case) or supraintuitionist (for the single-succedent case)—for base vocabularies that satisfy Containment. Further, the implications and incompatibilities that hold in virtue of logic alone, in the sense of holding no matter what base vocabulary the logically extended vocabulary is elaborated from, can be fully structurally closed: both monotonic and transitive. Of course, there are many implications and incompatibilities that hold in the logically extended language that *do* depend on the base vocabulary, and do not hold in virtue of the logical connective definitions alone. It is these we are particularly interested in. For it is these nonlogical reason relations holding in the logically extended vocabulary that make it possible to *say* there what the reason relations of the base vocabulary (and also its logical extension) are.

In addition to making explicit the material consequence and incompatibility relations that govern the use of arbitrary open-structured or substructural base vocabularies, expressivist logics also aspire to mark explicitly any local regions of those reason relations that do exhibit structural features that do not hold globally. Even material consequence and incompatibility relations that are not globally monotonic or transitive typically include some particular implications and incompatibilities that are persistent, that is that do hold monotonically. Some implications will satisfy Cumulative Transitivity. The regions of the reason relations where both monotonicity and transitivity hold amount to local islands of topological closure of reason relations within the larger sea of open-structured ones. Structural logicism about reason relations has obscured the intricate open structure of the implications and incompatibility that actually articulate the conceptual content of much material, nonlogical vocabulary. Expressing these features of the fine structure of material reason relations of base vocabularies in extensions of those vocabularies elaborated by introducing logical vocabulary is also something expressivists would like their logics to be able to do. Chapter Three shows how this expressivist criterion of adequacy can also be satisfied. Chapter Five introduces implication-space semantics, which is a model-theoretic metavocabulary with the expressive power to make explicit the fine structure of the conceptual roles that sentences of open-structured base vocabularies play in virtue of standing in reason relations of these intricate kinds.

Mentioning these two dimensions along which expressivist logics can hope to make reason relations explicit—saying *that* particular implications and incompatibilities hold and explicitly marking implications and incompatibilities that exhibit structural features that hold only locally and not globally within particular constellations of material reason relations—points in the direction of a response to a worry that residual structural logicist commitments and temptations might raise about the very idea of open-structured reason relations. For Tarski and Gentzen, and most of their heirs and admirers, the topological closure structure of logical consequence relations is essential to anything that deserves to be called a *consequence* relation. Similar remarks apply to the monotonicity of inconsistency, and even the explosion principle *ex contradictione quodlibet* concerning the interaction of implication and incompatibility relations. If, as we are recommending here, we drop *all* of those structural constraints on reason relations, what remains of the claim that they are *reason* relations? It might seem that we are just talking about relations, *tout court*. We are insisting that there are two kinds of reason relations, what we *call* "implication" and "incompatibility." The latter are symmetric, and the former are not. But what is it about a set of relations, some symmetric and some not, in virtue of which they deserve to count as "reason" relations? Why isn't this a case of throwing the baby of reasons out with the bathwater of closure structure? If structural logicists are wrong in appealing to structural features such as monotonicity and transitivity as essential features demarcating *rational* relations, what *does* pick out some relations among sentences as rational relations—distinguishing them from, say, orthographic or syntactic relations?

We have been suggesting that what is of most philosophical interest, and what logical expressivists should be most interested in being able explicitly to codify in the form of sentences that can themselves stand in reason relations, is the fine structure of the material reason relations exhibited by particular nonlogical vocabularies, such as regions of subjunctive robustness (local persistence of implication or incompatibility) and behavior under explicitation of implicit consequences. These intricate, detailed rational relations among sets of sentences only become visible once the traditional global structural demands that amount to requiring topological closure are relaxed. But if all that remains at the level of grosser global algebraic structure is so thin and abstract, doesn't that undercut the claim that what we are talking about is still relations of implication and incompatibility? What does it mean for the relations we are discussing to *be* relations of implication and incompatibility, what we have been calling "reason relations"?

We take this challenge very seriously. One of the principal tasks of this book is the development of an extended response to it. As we say in the Introduction, the form of that response is a kind of functionalism about reasons at the metatheoretic level. What makes the open-structured relations we discuss qualify or deserve to count as *reason* relations of implication and incompatibility is the role they play with respect to four kinds of metalanguage. We saw in Chapter One how relations of implication and incompatibility can be understood in terms of the use of sentences to make claims and defend and challenge them, when those discursive practices are specified in a bilateral normative *pragmatic* metavocabulary. That account neither entails nor presupposes that those relations have an algebraic or topological closure structure. Chapter Three shows how even the radically open-structured relations we are calling "implication" and "incompatibility" can be codified in a *logical* metavocabulary in the form of a familiar sort of multisuccedent sequent calculus that is elaborated from and explicative of arbitrary material base vocabularies, regardless of whether their structure is open or closed in the various senses we have distinguished. Chapter Four shows how to understand implication and incompatibility in a Finean truth-maker *semantic* metavocabulary from which all requirements of closure structure have been removed. It further reveals a remarkable isomorphism between that truth-maker semantic specification of reason relations and the bilateral normative pragmatic specification of them. Chapter Five then offers a metavocabulary for defining and manipulating the *conceptual roles* (showing up there as "implicational roles") sentences can play in virtue of the open-structured reason relations they stand in to one another. Our *metarational functionalist* claim is that a good thing to mean by "reason relations" is whatever plays all four of these roles: pragmatic, logical, semantic, and in the specification of conceptual roles, regardless of how spare and abstract the gross global structural constraints imposed on those relations might be. If it walks like a duck (pragmatically), quacks like a duck (in expressive logics), looks like a duck (semantically), and lays eggs (conceptual roles) like a duck, one is entitled to call it a duck.

This metalevel functionalism about reason relations also highlights a substantial difference between the very abstract account of reasons presented here and substantive contemporary accounts of rationality such as Bayesianism and rational choice theory. Given the prominence and popularity of such programs, our focus here on logicism as a way of thinking about reasons, and so rationality, might have seemed odd, blinkered, or perhaps merely old-fashioned. After all, for many contemporary thinkers, classical logic only matters insofar as it is a limiting case of probability theory, resulting from allowing probabilities to take only the values 1 and 0. These theories indeed formally define clear and precise ways of understanding the concept of good reason. But they are engaged in a very different sort of enterprise than the one pursued here. This is clear from the fact that the sentences to which credences are attached are taken to be already conceptually contentful, in advance of playing the role in reasoning Bayesians reconstruct. In rational choice theory, the sentences specifying options and outcomes must already be semantically interpreted in order to support assignments of conditional probability of outcomes relative to options, and preference and utility orderings on outcomes. The sense of "good reason" these theories define is accordingly not a candidate for explaining the semantic contentfulness of the sentences whose role in good reasoning is being defined. The semantics is outsourced, or rather just presupposed as available when the Bayesian or rational choice theorist comes on the scene. In this sense, our spare, highly abstract conception of reason relations aims to be explanatorily more basic than and conceptually prior to the late-coming substantive conceptions of rationality and good reason these approaches articulate.

This orienting aspiration—to specify a conception of reason relations that is capable of underwriting an understanding of the conceptual contents of sentences as roles they play in reason relations—might well motivate a skepticism about our enterprise based on the liberality of our abstract specification of reason relations, rather than skepticism based on commitment to structural logicism. For we are thinking to begin with of what stands in reason relations as just sentences as lexical items (what Sellars calls "sign designs"), not as the conceptual contents or propositions they express. How, one might wonder, can relations among such initially meaningless items (corresponding to Wittgenstein's "sign-post considered just as a piece of wood") be intelligible as *reason* relations of implication and incompatibility? What could justify calling them that? The programmatic commitment in question is an essential feature of the pragmatics-first order of explanation that seeks to explain how discursive practices of using sentences to make claims and challenge and defend them by making further claims that thereby have the practical significance for other practitioners of serving as reasons for or against them can confer conceptual contents on those sentences. We defend it by specifying the multidimensional metatheoretic role we claim characterizes reason relations as such. We have already offered a pragmatic story about the role in such practices relations among sentences must play in order to qualify as reason relations of implication and incompatibility. It will be supplemented in the chapters to come by an account of how reason relations in that pragmatic sense can be expressed explicitly in logical metavocabularies, using conditionals to express implications and negation (with conditionals) to express incompatibilities, as well as using modal operators to express local regions of structure. Abstract, radically substructural material reason relations then show up from a different metatheoretical perspective as interpretable in robust model-theoretic truth-maker semantic metavocabularies that can be proven to pick out the very same relations of implication and incompatibility previously picked out pragmatically and expressed logically. Formally tractable and usefully manipulable conceptual roles fit for substantial further explanatory work turn out to be definable from those very same abstractly characterized open-structured relations. In the next three chapters, in the course of elaborating specially constructed versions of these different kinds of discursive metavocabulary, we demonstrate a web of systematic relations among them that together define a precise sense in which they provide perspectives on a common topic: reason relations of implication and incompatibility.

Notes

1 In this work we do not consider *practical* reasons, or reasons for commitments other than specifically doxastic ones.
2 This idiom is introduced in Brandom (2008), Chapter Two.
3 See Belnap (1962). Complications to this analysis ensue in substructural settings—in particular, nontransitive ones, where Cut fails. See Ripley (2015).
4 This line of thought is presented in more detail in Chapter One of Brandom (2000), and Chapter Two of Brandom (1994).
5 From our point of view, "negation on the Australian plan" is just logical expressivism considering only one bit of logical vocabulary: negation.
6 Tarski put forward a model-theoretic version of this approach that appeals to permutation rather than substitution, which has been further developed and championed in our own time by Gila Sher (1991). At a very abstract level, the difference between these is whether we look at variations *salva consequentia* or *salva veritate*.
7 https://protect-us.mimecast.com/s/ZfbuCADQm5Ur4XPZjU9Ttvg?domain=aeon.co
8 As explained below, their projects of building nonmonotonic logics out of classical monotonic ones is alien to the expressivist approach. Expressivists want logics that are expressively adequate to codify nonmonotonic reason relations of implication and incompatibility. That is not at all the same enterprise as constructing *logics* that are nonmonotonic. In fact it turns out that the purely logical reason relations governing such logics (the implications and incompatibilities that hold in virtue of logic alone) governing logics that codify nonmonotonic material consequence relations can themselves be monotonic and transitive: structurally closed. So in the end we are in a position to justify the invocation of classical negation and inconsistency as part of the definition of one's nonmonotonic logic—if one still wants to do what these nonmonotonic logicians want to do.
9 We discuss the codification in implication-space semantics of reason relations of this sort near the end of Chapter Five, in Section 5.5.4.
10 This case has been made forcefully by Dov Gabbay (1985), who includes also CO and CT as necessary for workable nonmonotonic systems.
11 "A version of" because CT is additive, that is, context sharing, while Gentzen's Cut was multiplicative, that is, context combining. In open, substructural settings, these diverge.
12 This is a distinct, new notion of explicitation, not to be confused with the one that is involved in LX-ness. It does not specifically involve logical vocabulary. In Chapters Three and Five we will introduce still further layered senses to this multifarious concept, which is at the center of our expressivism.
13 We do not discuss the other most popular candidate weaker than MO, often called "rational monotony," because as usually formulated, it assumes the language already has negation in it. It says that no conclusions of a premise-set are lost by adding new premises that do not contradict it. There is a version that appeals only to incompatibility, but looking at structural principles relating

implication and incompatibility would take us too far beyond the argument here.

14 CO comes from Tarski's plausible condition that $X \subseteq Cn(X)$: any premise-set is contained in its consequence set. In the context of MO, CO is equivalent to Reflexivity (RE), which is what Gentzen actually uses (all leaves of all purely logical sequent derivations are instances of RE). If we relax MO, this equivalence breaks down, and one might worry, as relevance logicians do, about endorsing even the weak sort of monotonicity that CO enforces: monotonicity of implications of the form $A \mathrel{|\!\sim} A$, which can be arbitrarily weakened with further premises. But CO remains plausible when thought of in explicitation terms: what a premise-set *explicitly* contains counts as also *implicitly* contained in it. Explicit content is part of implicit content.

15 To jump ahead in our story a bit: the Incoherence-Incompatibility (II) codifying definition of negation introduced earlier entails that an incoherent set implies the negations of all of its members. If we explicitate those implications, add any of those negations as explicit further premises, the result is a premise-set that for some sentences contains both A and $\neg A$. Such premise-sets are *persistently* explicitly incoherent. If implicational explosion is retained for persistently incoherent sets, then taking the explicative path through these logically complex negated sentences will close off the possibility of further exploration by explicitation. The capacity to make reason relations explicit is a two-edged sword. With great expressive power comes great hermeneutic responsibility.