

abilities, and the various corresponding modes of rational integration afforded by them. Robert Brandom distinguishes the structural levels of semantic complexity of concepts afforded by the pragmatic interface between syntax and semantic as follows:

- concepts that only label and concepts that describe,
- ingredient and freestanding conceptual contents, making explicit the distinction between the content of concepts and the force of applying them, and
- concepts expressible already by simple predicates and concepts expressible only by complex predicates.⁶⁰

In its specialized format as the infrastructure of all *theoretical* structurations of the world, language is essentially formal. The significance of formal languages lies in their capacity for what Catarina Dutilh Novaes terms ‘desemantification and resemantification’.⁶¹ This is the ability of formal languages to be detached and abstracted from any particular content (topic-independence) so as to be generally applied to different contexts. The desemantifying ability of formal languages is tantamount to the explicit re-enactment of mind outside of any particular individual experience or contextual meaning—formal languages as the prostheses of extended cognition and epistemic enablement. Decoupled from any particular content, ‘desemantification allows for the deployment of reasoning strategies other than our default strategies, thus enhancing the “mind-altering” effect of reasoning with formalisms’.⁶² In addition, formal syntactic languages can be explicitly computable. This is the achievement of Noam Chomsky, in putting forward a hierarchy of formal grammar or syntax where the complexity of syntax is articulated in terms of computational complexity. Chomsky’s hierarchy classifies different types of syntax (recursively enumerable, context-sensitive, context-free and regular)

60 Brandom, *Reason in Philosophy*, 199.

61 See C. Dutilh Novaes, *Formal Languages in Logic* (Cambridge: Cambridge University Press, 2012).

62 Ibid., 219.

in terms of the combinations of generative processes required to produce them, the formal grammatical properties that specify them, and the automata necessary for computing them.⁶³

Finally, in its ultimate superior mode, language refers to artificial general languages—languages in which the full hierarchy of syntactic complexity (of formal languages) and the full hierarchy of semantic complexity (of natural language), computation and conception, desementification, resemantification, and semantic enrichment all exist side by side, reinforcing one another. Before proceeding, let us briefly look at the phenomenon of language—although the topic of language and its role in the formation of minding activities (judgements and inferences) will be surveyed in more detail in chapters 6 and 7.

The emphasis on the principal role of language for mind is an old topic. But what has been largely absent in studying the role of language for mind is the deep logico-computational structure of language itself, where language can be seen as much as a multilevel syntactic complexity as a multilevel semantic complexity. This is language neither as a medium of direct access to reality nor as a system of public discourse, but as a framework of interaction-as-computation incorporating different classes of semantic and syntactic complexity and the cognitive-practical abilities associated with them. This interactive computation permits qualitative compression of data and selectivity of compression, it significantly reduces the size of the agent's internal model while increasing its complexity, and it can format and modulate the agent's behaviours, stabilizing the multi-agent epistemic dynamic without which it is impossible for any agent to be aware of itself and to experience. Built on this computational interactive dimension, language is above all an engine for the generation of qualitatively distinct cognitive abilities. It is precisely what reshapes intelligent behaviours not by degree but in type (e.g., a linguistic agent differs from a nonlinguistic agent in type, not in degree).

63 See N. Chomsky, *Aspects of the Theory of Syntax* (Cambridge, MA: MIT Press, 1965); and for more details, R. Hausser, *Foundations of Computational Linguistics: Human-Computer Communication in Natural Language* (Dordrecht: Springer, 1999).

The archetypal figure behind language is computational duality or interaction. Interactions are the prevalent phenomena in language, and they can be concurrent, synchronic, asynchronic, typed or untyped, deterministic or nondeterministic. Interactive systems are online, open, and reactive to multiple streams of input. Interactions are governed by computational dualities. In simple terms, these dualities signify the interchange of roles between two or more processes or behaviours which constrain one another, resulting in the generation of additional constraints or rules which increase the complexity of behaviours involved in interaction. In theoretical computer science, these interactive dualities are called 'open harnesses'. Open harnesses simultaneously constrain the behaviours of the interacting systems (hence the constraining connotation of 'harness'), and harness them to a new design (a behaviour with a higher level of complexity). By pitting two systems against one another, they force the systems to correct their behaviours and to augment their capacities, as in a game that harnesses one player to respond intelligently to another, and where the totality of the game is always in excess of its players. But this is not a game in the game-theoretical sense: as topics of logic and theoretical computer science, interaction games are devoid of precise goals, payoff functions, or predetermined winning strategies or procedural rules for how the game should be played. Instead, the rules of the game emerge naturally from the interaction itself. Interactions are a medium of complexification for processes. In the words of Jean-Baptiste Joinet:

As vague as the notion of process may be, it certainly is in their nature not only to evolve (amongst the effects produced by a process appear its own transformations), but also to produce effects on other processes, mutual effects. In short, processes act and interact. Whatever its technological matter, whatever its implementation, the essence of a process is completely involved in its (potential) dynamic behavior: not only its own possible destinies under evaluation, but also the full set of possible operational effects it will occasion in all possible processes' interaction contexts. With respect to semantics, the answer brought by processes is

thus of a radically new kind, which implements the performative way of meaning: doing, is the way processes speak.⁶⁴

But unlike generic interactive processes, the logico-computational interactions of language have the ability to increasingly incorporate lower-level interactions within higher-level interactions (i.e., interactions with more semantic complexity or, in Wolfendale's terms, interactions that reformat lower-level interactions) and to form stable linguistic items with their own specific transformation rules across different levels. It is in the context of a complex network of interaction that rules are obtained and stabilized; some may be context-sensitive, others may be applied across the board. Semantic values or meanings are abstractions of linguistic locutions together with transformation rules for their use, which are both derived and used in their stabilized form through the interaction.

The realization of mind or general intelligence is inconceivable without language not merely as a structuring edifice but also as a necessary and vast computational framework for the generation of higher-order cognitive faculties (theoretical and practical judgements). Just as there is no structure in the world without the structuring mind, there is no mind and no *unrestricted* world without the structuration of language and its unrestricted universe of discourse wherein everything can be questioned or subjected to systematic theorization. General language (as opposed to this or that language) has no borders or limits. In the same way that we cannot step outside of mind to gain direct access to reality, we cannot step outside of language as the armature of mind. In stepping outside of one language, we only find ourselves in another more general language:

[T]o speak of a limit drawn by language (or by a language) is to be, linguistically, beyond the putative limit. A limit is a limit only if there is something beyond it, so to identify a limit to language is also to enter the linguistic space making it possible to speak of what is beyond the

64 J.-B. Joinet, 'Proofs, Reasoning and the Metamorphosis of Logic', in L.C. Pereira, E. Haeusler and V. de Paiva, *Advances in Natural Deduction* (Dordrecht: Springer, 2014), 58.

limit, and thus to negate the identification of the 'limit' as a limit of the language that bespeaks it as a limit.⁶⁵

The phenomenon of language in its full syntactic and semantic complexity cannot be captured outside of the logico-computational linguistic interactions of which pragmatics (meaning as use or discursive meaningfulness) is only the uppermost visible instantiation. Since mind or general intelligence is inconceivable without language, and to the extent that language requires an interactive framework, the idea that mind or general intelligence could be realized in anything but a community of agents is a dubious one. Humans are only minded and minding agents in that they have sociality. In the same vein, the idea of an artificial realization of general intelligence in anything but a multi-agent system rather than a single agent is something of a bygone twentieth-century science fiction. Artificial general intelligence is a product of interactions, be they between nurturing humans and child machines, or between machines that have graduated into the domain of artificial general languages.

To this end, general intelligence should be seen not only as a repertoire of existing cognitive abilities but also as a generative framework for the realization of new cognitive abilities by adjusting to the syntactic-semantic resources of language. It is the linguistically charged competence to proliferate, diversify, and maximize theoretical and practical abilities that sets general intelligence apart from complex causal and pattern-governed processes exhibiting a powerful yet restricted range of behaviours. The intrinsic affinity between general intelligence and language as a socially embedded and constructive medium for the elaboration and realization of the abilities of mind points to the social bases of general intelligence. There is no predetermined limit to the type and range of cognitive technologies that can be garnered by excavating 'the visible invisible essence' of geist or language. The possibilities of what can be done with language are as unfathomed as the possibilities of what language can do to its users.

65 Puntel, *Structure and Being*, 31.

**THE NECESSARY AND SUFFICIENT LINKS
BETWEEN CONCEPTION AND TRANSFORMATION**

Having presented an overview of mind not as an ideal object but as a social project, we can now shift the focus to the ramifications of developing and partaking in such a project.

Hegel's systematic elaboration of *geist* in functional terms has both a critical and a constructive import. The functional analysis of the spiritual or *geistig* struggle—i.e., the struggle of spirit to become the object of its own concept, to refine the content of this concept, and to devise abilities to actualize itself as the object of its concept—reveals how this struggle is realized, and how it can degenerate and precipitate pathologies of consciousness and their social manifestations. This type of analysis sets the stage for an in-depth diagnosis through which it is possible to identify what ought to be changed, and to specify those structural joints or material organizations that ought to be treated as sites of struggle. It is in this sense that Hegel's deep functional picture of *geist* as that which has a history and has the ability to treat its history as a milieu of intelligibility in which all achieved totalities are fleeting shadows and appearances lays the groundwork for Marx's project of communism as the *real movement* that concretely and determinately negates the present state of affairs, the givens of history. The uncovering of history as a new domain of intelligibilities requiring special kinds of theoretical and practical cognitions brings to light those sites where struggle and intervention must take place in order to achieve a consequential and concrete liberation from the present state of affairs perceived as the totality of history.

The import of the functional picture of *geist* resides in how this picture highlights key activities and abilities through which the mind as a social project can be augmented and amplified. Primary among these activities are conception and transformation; both are labours of mind, but one places more of an accent on systematic theorization, the other tends toward concrete practice. However, at no time can these two be separated: they are as mutually reinforcing as they are intrinsically linked.

Essential self-consciousness is explicated by the necessary link between self-conception and self-transformation as two distinct yet connected

activities: that of conceiving (i.e., bringing into conception) ourselves in the world, and that of transforming ourselves in accordance with our self-conceptions.⁶⁶ Each transformation serves in turn as a pivot for a different encounter with ourselves in the world. Alteration in the order of self-conception induces a change in the self of which it is a conception, and a qualitative difference in the self induces a transformation in the general or geistig conception of the self. The positive feedback loop between conception and transformation instigates a disequilibrium, with each difference in one order destabilizing the other and becoming a cue for its readaptation. This disequilibrium is particularly intensified as subsequent conceptions become further removed from the order of appearances, and in the process the latter are determinately negated. The act of self-conception must yield progressively more fine-grained conceptions of the geistig self (what it means to be a minded and minding agent, a human, an embodiment of intelligence); but in order to do so, it must break away from the mere appearances of its present conception, its manifest realization.

In conceiving itself in terms of this process of simultaneous self-recognition and self-negation, the history of geist is no longer *entailed* by its past states, even if it is constructed out of them. The self-conception of geist via determinate and concrete negation of its conceptual content (how it appears to itself) brings about a real movement of history in which geist's sequence of self-constituted transformations are not merely the repetitions of its past stages. Within this movement, any totalized state of realization is taken to be a historical appearance with regard to the reality of history qua incomplete and ongoing totalization. The history of geist, in this sense, is neither a linear progression between different states of geist's realization (epochs) nor a series of isomorphic transformations. The reality of what geist is now is not the reality that it used to be. What geist or intelligence will be is never what it is or has been, where 'never' implies the elimination of the one-to-one relationship between the sequences of geist's transformation. Historical consciousness, or awareness of this fact—what

66 See R. Brandom, *A Spirit of Trust* (2014), <http://www.pitt.edu/~brandom/spirit_of_trust_2014.html>.

it means to identify the appearances of history and determinately negate them—is the kernel of concrete self-consciousness as the real movement of emancipation.

The adequacy of self-consciousness is an index of the quality of conception and transformation and an indication of how effectively self-conception and self-transformation are integrated within the order of intelligibilities. Just as a robust self-conception brings about a consequential self-transformation, a flawed or impoverished self-conception incites a defective or inconsequential self-transformation. But in so far as our self-transformation is embedded in the objective world, its flaws are not isolated. Depending on its proportion and persistence, a pathological self-transformation can trigger a disaster by laying ruin to those structures that environ and support us. This calamity also debilitates our capacities to act, either simply by depleting the resources required for perpetuation and transformation, or more insidiously by causing a disorientation that prevents us from decisively thinking and taking measures, detecting and acting on opportunities without the crippling anxiety of further deterioration and imaginary tragedies. Blinded by the oppressive fear of tragedies, the diagnosis of what exactly is wrong (revealing the 'specific' inadequacies of conception *qua* determinate negation) then falls under the jurisdiction of abstract negation and unconscious self-deception. In supposing ourselves to have dispensed with all illusions and escaped all the traps of history, we succumb to naivety and cognitive-practical despondency.

In this environment, reasons which form and adjudicate recognitions, obligations, responsibilities, and pertinencies (what matters and what does not) are discredited by a critical discourse on the irrationality of reasons tout court, and an insistence that we should dispense with them. Reasons are unmasked as social irrational causes that merely disguise the roots of what is oppressive and exploitative. This critical discourse then becomes an exposé on how causes not only semantically distort reasons but also masquerade as reasons. Reason is burnt at the stake, accused of being the ultimate trap of history, the supreme collaborator with the oppressor, a smokescreen for the conditions of exploitation. Once freed from the tyranny of reason, critical discourse presents itself as an egalitarian exposition of

the causes of oppression and exploitation. Having immolated the trap of all traps—reason—critique is now fully disillusioned and emancipated. But to dispense with the conceptual and normative resources of reason in order to expose their causal irrationality, and to diagnose oppressive pathologies via the reasonless appraisal of causes, amounts to the impoverishment of the semantic requirements necessary for the intelligibility of any diagnosis. In other words, the unconscious undermining of the criteria required for the adequate consciousness of the self occasions a consciousness which is inadequate to have a proper conception of itself in the world and to determinately transform itself in accordance with such a conception. The reasonless critique or diagnosis of what is oppressive and exploitative becomes the prognostic course of the disease or pathology itself. In being an inadequate consciousness, the critique unconsciously justifies those pathologies of history it is bent on eradicating.⁶⁷

Founded on the intelligibility of the concept and the intelligibility of practices, the adequate self-consciousness is distinguished as a multitasking project comprised of four basic undertakings:

- (1) To evaluate and correct our self-conceptions (using the resources of both common-sense rationality and those of the sciences, with the understanding that they yield two different images of ourselves in the world that should be integrated).
- (2) To coordinate our self-transformations with better self-conceptions (i.e., self-conceptions informed by a broader order of intelligibilities).
- (3) To amplify the influence of our rational and scientifically informed self-conceptions over our self-transformations (the shift from a *necessary* link between the two to one of *sufficiency*).

67 See R. Brandom, *Reason, Genealogy, and the Hermeneutics of Magnanimity* (2014), <<http://www.pitt.edu/~brandom/downloads/RGHM%20%2012-11-21%20a.docx>>; and Brassier, 'Dialectics Between Suspicion and Trust', 98–113.

- (4) To revise and where necessary abandon our self-conceptions in accordance with the intelligibility of *how* and *into what* we are transformed.

Formulating self-consciousness as a project rather than as an ideal object allows us to conceive of the vocations of thought and action not on an ontological basis, where the ideal abstractness of thought is contrasted with the ideal concreteness of action, but on a methodological basis, where thought must first make a concrete difference in itself in order to make a difference in the world. The interplay between the abstract and the concrete is then conceived methodologically as the determination of concrete abstraction and the abstraction of the concrete—a dialectics through which making a difference in thought and making a difference in the world can be *bidirectionally* mediated by autonomous transformative actions in the order of thought and in the order of the world. Autonomy in reference to transformative actions describes the ability of these actions to repurpose or reorient themselves in accordance with self-contained rational ends. It is important to note that to speak of functional self-containment in this sense does not imply the denial of the existence of material constraints. In fact, it simultaneously brings about the opportunity to uncover such constraints and to examine how they can be modified.

GEIST AT THE EDGE OF INTELLIGIBILITY

To every abstract moment of science, there corresponds a shape of appearing spirit per se. Just as existing spirit is not richer than science, so too spirit in its content is no poorer.⁶⁸

The reinforcing link between self-conception and self-transformation characterises essentially self-conscious creatures or rational agents as those possessing the capacity to constitute a history and to have a contentful experience of it, to have impersonal norms regarding what ought or ought not to be done in order to maintain and expand the intelligibility of their history.

68 Hegel, *Phenomenology of Spirit*, §805.

This capacity unfolds the truth of intelligence as a social cognitive-practical enterprise in the order of intelligibilities, of things, practices and values. In reality, the systematic encounter with the order of intelligibilities is the inexorable consequence of the realization of geist by the specific qualitative set of activities that distinguish it.

The formulation of self-consciousness as a matter of practical achievement underlines the import of self-conception as a recipe for action without which action cannot maintain its practical intelligibility. Conception without praxis is unrealized abstraction and praxis without conception is a hollow impression of concreteness. Self-conception in its concrete form is a search for intelligibilities pertaining to the world of which we are part. Self-transformation in its consequential form is the intelligibility of practices in response to the intelligibility of what the world we inhabit is, together with what we ought to be in accordance with the ends of thought. The recognition of ourselves, the conceiving of ourselves within one or multiple self-narratives, is by definition the construction of ourselves within the order of intelligibilities. There could be no 'us' without our encountering ourselves and bringing this encounter into a conception not only of who or what we are, but also of where we have come from, where we are, which paths have led to where we are, and which paths we ought to take. The conception of this encounter is thus an open landscape of inquiry into different orders of intelligibilities. The conceptual awareness of one's experience of the world is a necessary framework within which one can know oneself in the world, and how to change oneself and the world.

However, the divergence between what we take ourselves to be and what we actually are, the disparity between the intelligibility of the self and the world as what they appear to be and the intelligibility of the world in itself as what conditions all appearances generates a tension in self-conception and, correspondingly, in the space of self-consciousness as a project. The full recognition of this tension—its simultaneous sharpening and resolution—characterizes the task of scientific rationality or, more generally, the enterprise of science. Again, what is meant by science here is not just the modern empirical sciences, but also the science of history, the science of thinking, or what Hegel dubs the Greater Logic,

and finally the science of impersonal values and disvalues, the science of ethics. While maintaining this broad idea of science, I would however like to discuss the implications of the modern empirical sciences for the project of conception and transformation outlined above. For the field of empirical sciences poses a significant challenge to what is already a challenging project.

The core protocol of modern empirical science consists in striving for explanation. Why do we see what appears—in our ordinary common-sense experience—to be the case, and yet know that it is not the whole of reality? Science overturns the order of what appears to be subjectively the case by providing its objective explanans. It differentiates what was previously an apparent explanation into an explanandum the explanans of which remains to be found. The explanatory force of science reaches its peak when it overturns the universe that seemed to be ontologically dependent on the mind into a universe that is independent of mind. Yet in overturning what is simply subjectively manifest—the universe hypostatized as the reificatory dependent object of mind—science draws on the core components of subjective or minding *activities*: linguistic doings, conceptualization, and systematic theorization among others. Objective science without subject-constituting mind is a subjective delusion. The movement of scientific inquiry only underlines the necessity of mind in its pure and necessary form. Science, in this sense, is the hallmark of a mind that has matured to learn that it should hypostatize neither its structuring activities nor its structured object, neither itself nor its unrestricted universe. This is a mind whose intelligibility is not immediately given in itself, but is achievable only in its integration—rather than complete fusion or reduction—with an intelligible order pertaining to a non-manifest and mind-independent reality. The enterprise of science is permanently caught up within this integral framework which is comprised of the vocabularies of mind and the items of the world, and whose internal tensions enrich and drive it. But this tension through which science expands itself is exactly the tension exclusive and constitutive to mind, albeit more self-conscious and intensified.

Spirit knowing itself in that way as spirit is science. Science is its actuality, and science is the realm it builds for itself in its own proper element.⁶⁹

With the advent of scientific explanation and the excavation of an intelligible universe which, in its objectivity, constrains our thoughts about it, the order of self-consciousness—the duality of the intelligible and intelligence—enters a new phase. Matter-of-factual truths pertaining to a nonmanifest and mind-independent universe intrude upon and are ingrained in the logico-conceptual infrastructure of intelligence. As the scientific will-to-explain or the pursuit of explanatory coherence highlights the insufficiency of common-sense concepts, conception begins to adapt itself to science and its unrestricted universe.

Even though the conceptual framework of common sense through which we articulate the intelligibility of ourselves in the world is built upon basic matter-of-factual truths, it does not grant us access to these truths as such; properly conceived, since they lie beyond what manifestly appears to be the case, these truths belong to the framework of scientific inquiry. By assuming otherwise we subscribe to the ideological fixation of the epistemic given—of access to the intelligibility of the world through the supposedly spontaneous intelligibility of our seemings.

Scientific explanation, accordingly, can be understood as an activity of mind that forces the activity of self-conception to renegotiate its original alliance with appearances—particularly the appearance of how the world that appears for mind is in itself. As the order of intelligibilities uncovered by science encroaches upon the intelligibility of that which is manifest, self-consciousness gains a new mobility. The sharp asymmetry between what is manifest and what is scientific, how things appear to be from an ordinary common-sense perspective and how things are as matter of fact, destabilizes the oppressive serenity of the order of mind and things. Cognitive progress can only be maintained and expanded by sharpening this asymmetry, by further amplifying the instability in how mind and the world stand in relation to one another. As the site of this tension, science

69 Ibid., §25.

advances through positing the transcendental excess of a structuring mind and by positing a structured reality that is in excess of mind.⁷⁰

Yet the advancement of science need not be understood as a series approaching convergence, or a uniform progression.⁷¹ A completed science, or the full and adequate scientific image (of human being in the world) is a regulative ideal. But, as Kant reminds us, when regulative ideals are applied constitutively they lead to contradictions.⁷² Outside of their regulative use, regulative ideas are pseudo-rational and illusory existences. The account of science as a convergent and uniform progression is not necessary for the certification of science as a vector of cognitive progress. Indeed, this convergent progressive interpretation can result in both irrationality and relativism, both epistemic dogmatism and epistemic anarchy. Wolfgang Stegmüller has provided an incisive critique of this generalized progressivist interpretation of scientific theories.⁷³ Here I will merely summarize his particularly meticulous and technical argument.

The idea of a general convergent progress of science is based on the analysis of the structure and dynamics of scientific theories (the theory-ladenness of all sciences). According to Stegmüller, the conviction of a convergent and uniform progress of science rests on the idea of the *general* reducibility of one theory (a dislodging theory) to another preceding (dislodged) theory covering the same class of observations. For example, if we can in every respect reduce or map statistical thermodynamics to the

70 'At its debut, where science has been brought neither to completeness of detail nor to perfection of form, it is open to reproach. However, even if it is unjust to suppose that this reproach even touches on the essence of science, it would be equally unjust and inadmissible not to honor the demand for the further development of science. This opposition seems to be the principal knot which scientific culture at present is struggling to loosen and which it does not yet properly understand.' Ibid., §14.

71 See for example, J. Rosenberg, *Wilfrid Sellars: Fusing the Images* (Oxford: Oxford University Press, 2007).

72 See I. Kant, *Critique of Pure Reason*, tr. P. Guyer and A.W. Wood (Cambridge: Cambridge University Press, 2000), 591

73 W. Stegmüller, *The Structure and Dynamics of Theories* (New York: Springer, 1976).

thermal theory of thermodynamics, then we can say there is a uniform or convergent progress from the latter to the former, from the less explanatory to the more explanatory theory. In the traditional study of scientific theories, analysis is performed at the micrological level, pertaining to the stable core of theories. These stable cores are comprised of atomic axiomatic formal sentences or classes of sentences together with their inferential relations. At the level of micrological analysis, to see whether one theory is reducible to another theory a mapping must be obtained from the stable core of the succeeding theory (T_2) to that of the preceding theory (T_1). But this mapping is precisely based on the reduction of individual theories to their so-called stable cores, a reduction that loses much necessary information regarding the specificity of its content, including 'the distinction between theoretical and nontheoretical functions, the general and special constraints, and the special laws holding only for certain intended applications'.⁷⁴ Accordingly, at the level of micrological analysis—pertaining to the stable cores or sentences and their inferential relations—the reducibility of T_2 to T_1 cannot be obtained in a nonarbitrary manner. In so far as micrological reducibility presupposes the prior reduction of each theory to its stable core, we cannot *generalize* the relationship between T_2 and T_1 in terms of their being generally more inductively simple and generally less inductively simple, generally more explanatory and generally less explanatory, answering more well-formed questions and answering less well-formed questions. The comparison of theory-contents is completely context-dependent. What appears to be less explanatory in T_2 may be less explanatory in one particular context than T_1 . Equally, T_1 may answer some well-formed questions that cannot be answered in T_2 , as in the case of Newton's and Einstein's theories of gravitation.

Just because T_1 precedes T_2 and both cover the same class of observations, it cannot be inferred that the relation between them is that of theory-reducibility in the sense mentioned above. It is only when we move from the level of micrological analysis to the level of macrological analysis, from the stable cores to the model-theoretic view of theories pertaining to their

74 Ibid., 127.

expanded cores, that we can speak of reducibility. The expanded core of each theory covers the class of partial possible models (i.e., the physical systems about which the theory is talking) or the class of sets of possible applications. Further, the components of the expanded core (its range of possible applications) are unstable and dynamically changing. To secure reducibility at this level of analysis, one has to take a static snapshot of theories at time t in order to obtain the mapping between one theory and another. But this static theory-comparison is itself problematic, since it freezes the dynamic picture of the theory structures and narrows the range of their possible applicability or coverage over possible models.

At the level of the micrological analysis of stable cores concerning formally constructed and inferentially related sentences or classes of sentences, the reducibility fails to be nonarbitrary or general. And at the level of the macrological analysis of expanded cores, reducibility can only be obtained by taking a static view of the dynamic structure of compared theories, hence losing information regarding the dynamic aspects of the theory structures themselves. Stegmüller's final conclusion is that we do not need a convergent progressivist account of scientific theories to talk about the rationality of science or to avoid relativism. Scientific progress can be seen in terms of the dynamics of each theory structure in itself—the increase in the range of its applicability and achievements as it incorporates aspects of an older theory structure which is now operating under new constraints and laws, and over a larger field of experience. Both the dislodged theory and the dislodging theory may prove successful in certain situations, identification of which requires a rich dynamic picture of their structures.

The true revolutionary import of science lies in its capacity to amplify reason's own power of knowing and to instigate cognitive expansion. The convergent progressivist interpretation of scientific theories is often assumed to be a preventive measure against irrationality and relativism with regard to scientific theories; but in reality it is an oversimplification that causes more unnecessary problems which can become the source of irrationality. The rationality of science lies not in the uniform progression of science, but in how these theories are constructed and how they expand the capacities of reason and its cognitive traction on the world. In light

of the dynamics of scientific structuration or theoreticity, the rationality of science can be preserved without a convergent progressivist reading of scientific change. Similarly, epistemological anarchism can be shown to be merely a parasitic outcome of the pseudo-rationality of an uncritically progressivist view of science.

Coming back to the role of science for mind, intelligence, and agency, science's capacity to make the world intelligible beyond appearances as a reality that is in excess of mind represents mind's movement of self-consciousness as it differentiates its universal and necessary features from its particular and contingent characteristics. The ongoing instability or perturbation caused by the sharp asymmetry between the manifest and scientific frameworks should be seen as a positive condition for self-conception and the corresponding self-transformation, since it results in the increasing refinement of the manifest picture of the agency, its minimalization to a set of logically-irreducible and necessary activities. Science is not an attack upon the logic or essence of the human that should be staved off; it is that which differentiates the necessary aspects of the latter from their contingent features. Pruning the manifest picture of the human and cutting it down to its logico-conceptual, necessary activities and functions is a required step toward understanding the meaning of the human. It is precisely through this minimalization of the manifest that the rational self can be extricated from the neurobiologically fabricated 'phenomenal self-model',⁷⁵ and instead presented as a constructive principle that can be transferred and artificially implemented. It is the continuous labour of science in deepening the order of intelligibility that provides geist with the necessary resources for the determinate negation and reappraisal of the *content* of self-conception.

In the wake of scientific rationality, mind turns into a wave of noetic deracination. This deracination of thought and its noetic drift is commensurate with what Plato calls the Form of Good as the Form of Forms, since it sets up the scaffolding for a conception of the realm of intelligibilities as

75 T. Metzinger, *Being No One: The Self-Model Theory of Subjectivity* (Cambridge, MA: MIT Press, 2003).

a complex system of recipes for crafting a world which includes not only satisfying lives but also the perpetual demand for the better. The ingredient of these recipes are not just theoretical intelligibilities, the products of modern sciences, but also practical intelligibilities and axiological intelligibilities, the objects of sciences of skill, practice, and ethics, all of which are subsets of the logical functions of mind. This is ethics as the science of impersonal values and disvalues concerning the intelligible unity of freedom of mind and the constraints of the world, the autonomy of thinking and reality in its otherness. But what kind of life would really satisfy the mind, 'other than one that involves a self-knowledge which has passed through all the stages of disciplined reflection on the source of things',⁷⁶ that is to say, their intelligibility? And what is intelligence other than that which knows what to do with the intelligible, whether pertaining to itself or to the world?

76 W. Sellars, *Essays in Philosophy and its History* (Dordrecht: D. Reidel, 1974), 26.

2. An Outside View of Ourselves as Experimental AGI (Problems, Concepts, and Models)

The aim of this and the following chapters is to engage—in a more detailed fashion—a question that overshadowed chapter 1's outline of the nature of mind as a configuring or structuring factor and as that which is capable of treating itself as an artefact of its own concept. This is the question of language as the *dasein* of *geist*—of language as a framework through which intelligence comes into cognitive contact with itself in the world. The question might be formulated as follows: What is it in language that makes the self-conscious form of intelligence not only possible but also amenable to self-determination and self-augmentation? To adequately answer this question, we have to reconstruct and explain the essence of language rather than simply highlighting its importance. If the ineffability of general intelligence is to be overcome, and if *geist*'s activities are fundamentally caught up in language, then we first of all need to understand language not in terms of some mysterious internal essence, but in terms of its computational capacities and formal autonomy; how it gains traction on the world and how it generates cognitive-practical abilities through which language-users can bring into conception—and potentially transform—themselves and their world. In other words, we need to inquire into *how language is realized* both at the level of natural evolution and that of social evolution, and *how language functions* both at its most fundamental level and at the level of what we might call the familiar picture of natural language and linguistic interaction (social discursive activities, the mappings between thoughts and speech acts, etc.).

To this end, our path to the nature of language, the nature of thinking, and general intelligence won't be anything like a straightforward nonstop train ride across different continents of inquiry. This is a route that stretches out with bends and twists from one condition necessary for the realization of mind to another. On this ride, language is the last station. At times we will stop the train to investigate the wilderness that we have only been able