Journal of Indian Council of Philosophical Research

is a tri-annual philosophical journal published by the Indian Council of Philosophical Research (ICPR). It is devoted to the publication of original papers of high standard in any branch of philosophy. One of the objectives of the ICPR is to encourage interdisciplinary research with direct philosophical relevance. Accordingly, contributions from scholars in other fields of knowledge, dealing with specific philosophical problems connected with their respective fields of specialization, would be highly welcome. However, good and original contributions pertaining to any branch of traditional philosophy would be equally welcome.

Each regular issue of the journal will contain, besides full-length papers, discussions and comments, notes on papers, book reviews, information on new books and other relevant academic information. Each issue will contain about 160–180 pages (Royal 8vo).

Annual Subscriptions

	Inland	Foreign	
Institutions	Rs. 300	U.S. $ 30	(Surface Mail)
Individuals	Rs. 150	U.S. $ 20	-do-
Students and retired teachers	Rs. 100	U.S. $ 15	-do-
Single Issue	Rs. 100	U.S. $ 15	-do-
Life Membership	Rs. 1200	U.S. $ 180	-do-

Bonafide students and retired teachers are requested to ask for special subscription form.

Air mail cost will be charged extra to those subscribers who want to get the journal by air mail. Request for air mail delivery must be made in writing.

For subscription and all other business enquiries (including advertisement in the JICPR) please contact directly:

Subscription Department
CENTRAL NEWS AGENCY PRIVATE LIMITED
23/90 Connaught Circus
New Delhi 110 001, India

All subscriptions must be prepaid.

All contributions to the Journal, other editorial enquiries and books for review are to be sent to the Editor, Indian Council of Philosophical Research, Rajendra Bhavan (Fourth Floor), 210 Deen Dayal Upadhyaya Marg, New Delhi 110 002.

Journal of Indian Council of Philosophical Research January–April 1998

ISSN 0970–7794

Journal of Indian Council of Philosophical Research

JICPR

Editor DAYA KRISHNA

Volume XV Number 2
January–April 1998

NOTE TO CONTRIBUTORS

JICPR welcomes contributions in all fields of Philosophy. However, it would like its contributors to focus on what they consider to be significantly new and important in what they have to say and to consider the counter arguments to what they are saying. This is to ensure that others may judge that what they are saying is on the whole more reasonable than the views opposed to their own. The historical preliminaries may be avoided unless they are absolutely necessary to the development of the argument, as it may be assumed that most of the readers of the *Journal* are already familiar with them. Reference and quotations are generally to be avoided except in an article that is specifically exegetical. Even in such cases the author is expected to give substantive reasons as to why he differs from the accepted interpretations. The article should, as far as possible, avoid jargon and the author's contention should be stated in as simple a language as possible.

The articles which use Sanskrit terms should use the standard diacritical marks, a specimen list of which is given at the end of the *Journal*.

Editorial correspondence, including manuscripts for submission should be sent to Prof. Daya Krishna, Editor, *Journal of the Indian Council of Philosophical Research*, B/189 A, University Marg, Bapu Nagar, Jaipur - 302 015, or to the Indian Council of Philosophical Research, Rajendra Bhavan, IVth Floor, 210, Deen Dayal Upadhyaya Marg, New Delhi, 110 002. Articles should be between 3000 to 6000 words, two copies of which should be submitted. They should be typed on one side of the paper, double spaced with ample margins, and contain author(s)/contributor's name and his/her institutional affiliation along with the complete mailing address. An abstract of 150–200 words should be included. Notes and references should appear at the end of the articles as *Notes*.

Only papers which have not been published elsewhere will be considered.

Proofs will be sent to the authors if there is sufficient time to do so. They should be corrected and returned to the *Journal* at the Delhi address given above within ten days. Major alterations to the text cannot be accepted.

Authors will be sent twenty-five off-prints of their articles free of charge. They may order more of the same on payment.

Articles in the *Journal* are indexed in the *Philosopher's Index*, USA.

STYLE SHEET

For the papers in English for the *Nirgrantha*
Transliteration Conventions

For the papers written in English, words from Sanskrit, Ardhamāgadhī and other Prakrits including the Apabhraṁśa etc., will be diacriticised if rendered in Roman script. (Quotations can also be in Nāgarī script.) (Here we suggest those for the Sanskrit (classical), the Prakrit, the Apabhraṁśa, and the Dravidic languages. For other languages, namely Arabic, Persian and the modern European languages, the current international conventions for transliteration for their rendering may be followed.)

Continued on back cover

JOURNAL OF INDIAN COUNCIL OF PHILOSOPHICAL RESEARCH

Volume XV
Number 2
January–April
1998

Editor: Daya Krishna

Indian Council of Philosophical Research
Rajendra Bhawan (Fourth Floor), Deen Dayal Upadhyaya Marg
New Delhi 110 002, India

Articles published in this journal are indexed in the
Philosophers' Index, USA.

Printed in India
at Chaman Offset Press, New Delhi 110 002
and published by Member-Secretary
for Indian Council of Philosophical Research
Rajendra Bhavan (Fourth Floor), 210 Deen Dayal Upadhyaya Marg
New Delhi 110 002

Contents

ARTICLES

SURVEY ARTICLE

DISCUSSION AND COMMENTS

Mill's Theory of Non-Provability

D. GUHA
Ranchi University, Ranchi 834 007, Bihar

The major objective of this paper is to show that the central theme of Mill's ethics is the *non-provability* of the ethical first principle, viz. 'Utility'. One who does not understand this basic point that Mill aims at elucidating in Chapter IV of 'Utilitarianism' fails to understand that Mill is a *rule utilitarian*, his ethics is an *ethics of principle* and that he is a *consistent hedonist*. Thus, one who fails to understand that Mill had in no sense logically proved the ethical first principle also suffers from three basic misconceptions, namely, Mill is an act utilitarian, his ethics is an ethics of virtue and that he is an inconsistent hedonist. In the course of this discussion, I shall, therefore, try to show how misconceived the interpretations and evaluation of some of the orthodox evaluators of Mill's ethics are and how important it is for us (the neo-evaluators) to dispel such misconception so that a proper understanding of Mill's theory of *justification* may be made possible. I call some of the evaluators orthodox in the sense that they mostly suffer from fixity of belief; viz., Mill is a supporter of provability thesis in ethics. This is to say that he either deductively or inductively proved that 'utility' ought to be first principle in morality. I call them orthodox in their approach because they would rather read Mill carelessly, suffer from some prejudice, give a literal interpretation of what he writes in his text and thus cling to some misconception, which they can very well avoid. On the other hand, neo-evaluators show a catholic spirit by reading Mill more carefully and *sympathetically*, refrain from prejudices and literal interpretation of his text and thus bring into limelight a few things erstwhile unknown or at least not very well known about Mill's ethics. One such thing is his non-provability theory and the allied notions mentioned earlier.

One who hurries through the epoch-making book of Mill's 'Utilitarianism' cannot overlook the title of the Chapter IV: 'Of what sort of proof the Principle of Utility is susceptible'. It is this chapter that attracted the minds of many orthodox evaluators of Mill's ethics, mainly Moore, Bradley and Sidgwick, in a negative way because they argued that Mill had in this chapter committed glaring mistakes both in proving logically (deductively and inductively) the first principle in morality, namely, 'Utility' and in proposing a strict logical definition of moral term

'good' in terms of a non-moral or natural expression 'the maximization of the general welfare'. I would attempt to show the hopelessness of such negative thinking and justify that Mill was in no sense different from the basic contention of at least Moore that no ethical first principle can be strictly logically proved within an ethical system and that an ethical term cannot be strictly logically defined by another non-moral or natural term.

I would like to mention an oft-quoted, oft-maligned and oft-misunderstood passage in Mill's 'Utilitarianism' (Ch. IV, p. 36) to understand Mill's position about proving clearly. He writes: 'The only proof capable of being given a thing is visible is that people actually see it. The only proof that a sound is audible is that people actually hear it; and so of the other sources of our experiences. In the like manner, I apprehend the sole evidence it is possible to produce that anything is desirable, is that people do actually desire it. If the end which the utilitarian doctrine proposes to itself were not, in theory and in practice, acknowledged to be an end, nothing could ever convince any person that it was so'.[1]

There are two reactions immediately, Bradley was 'ashamed' to have to examine such reasoning, 'but it is necessary to do so, since it is common enough'[2] and Moore reacts saying, 'Well, the fallacy in this step is so obvious that it is quite wonderful how Mill failed to see it'.[3] He gives this fallacy a name (that became famous)—the 'naturalistic fallacy'. He says, 'Mill has made naive and artless a use of the naturalistic fallacy as anybody would desire. "Good", he tells us, means "desirable" and you can only find out what is desirable by seeking to find out what is actually desired. The important step for ethics is this one just taken, the step which pretends to prove that "good" means desired'.[4]

If we go through Moore's view regarding the naturalistic fallacy, we get an idea that he had in mind some kind of a 'definist fallacy'. Prof. W.K. Frankena in an article on 'The naturalistic fallacy'[5] had tried to show that Moore's alleged imputation of this fallacy to Mill's ethics stems out of the idea that it is in fact a 'definist fallacy' because according to Moore, one who, like Mill, *defines* a moral term 'good' in terms of a natural term, 'happiness of the maximum', commits naturalistic fallacy. This confusion that Mill defines logically the moral term 'good' in terms of a natural term 'happiness of the maximum' leads the orthodox thinkers to believe that Mill holds a *definist view* of the justification of an ethical principle because they went on arguing that Mill's attempts at defining 'good' and 'utility' are a sheer result of *is-ought* confusion, i.e. 'ought' can be derived from 'is' (value in terms of fact). Thus, the naturalistic fallacy, when imputed to Mill's ethics, does not only make us believe that it is a kind of definist fallacy that Moore had in mind, he also had another orthodox belief, and that is that Mill necessarily, logically proves an ethical first principle from what *is*. Hence, Mill not only commits a definist fallacy, he also is a supporter of a definist theory of justification in ethics.

Frankena argues that this fallacy can occur only when a person (who commits this fallacy) distinguishes the properties of 'goodness' and 'happiness' in their ethical system. However, a *naturalist* who altogether denies any distinction between 'goodness' or 'desirableness' and 'happiness' or 'desirableness', may not be accused of having committed this fallacy at all. Hence, in the passage I have quoted from Mill's ethics, definist fallacy cannot be traced when Mill says, 'The desirable just is the desired'. This seems so obviously true for a psychological realist like Mill that even we wonder how Moore failed to grasp it. Moore even ran into an error in thinking that 'good' is indefinable, and to deny this necessarily involved a fallacy called 'the naturalistic fallacy'. The fact that a naturalist is one who need *not* necessarily commit naturalistic fallacy is so grossly overlooked by Moore that he forgets that the central teaching in 'Principia Ethica' is just this. This highlights an interesting fact and that is, that one may be like Mill and Moore, a naturalist, and yet not commit a blunder in defining a moral term in natural terms. A naturalist may be so called for quite different reasons. There is no point in calling an ethical thinker a naturalist who necessarily commits the above mentioned blunder. 'Who is a naturalist?', is a question to be answered later on. What we understand is that the definist fallacy is no fallacy unless the predicates definitionally identified are also taken to refer to different properties. So here, if Mill is saying that there is no property of desirableness, that it is consonant with usage to suppose that the word 'desirableness' just refers to desiredness, he has committed no fallacy whatsoever. Hence, like many neo-evaluators of Mill's ethics, I think that Moore's so called famous naturalistic fallacy relegates to infamy.

Moore's next misconception is about 'proving' an ethical first principle (due to the passage of Mill that I have quoted). Moore had correctly paraphrased to write, "You can only find out what is desirable by seeking to find out what is actually desired". But he seems to suddenly edify in air to say that Mill has pretended to *prove* that 'good' means 'desired'. This misconception of Moore that Mill had necessarily proved the desirable from what is desired, results in Moore's imputation of the naturalistic fallacy as a definist fallacy to Mill's ethics (and also resulted in a misconception that Mill is a supporter of definist theory of justification). But Mill was perfectly aware that 'desirable' does *not mean* 'able to be desired' and so, in this respect, was not at all analogous to 'visible'. 'Desirability' on safe grounds may be thought as *resembling* 'visibility' but it is *not* a safe ground to argue that, therefore, 'desirability' like 'visibility' is the 'desired' and 'visualized'. Moore has a peculiar obsession in his mind. He thinks that the occurrences of A with B is the same as saying that A is identified with B. If this be the case, then everyone who admits extensional equivalence of two properties shall commit extensionalist fallacy. Hence, Mill may not be accused of commiting any fallacy

pertaining to logically proving of the desirable first principle from actual desires or precisely, cannot be charged with deriving 'ought' from 'is'. Why strict logical proving cannot be charged against Mill can be understood only if we transmute the criticism of Moore in a syllogistic form of argument as:

The 'desirable' is the same as the 'desired',
The 'good' is the same as 'desirable',
∴ The 'good' is the same as the 'desired'.

What Moore has in mind is that in this syllogism, the middle term 'desirable' is ambiguous and this results in commission of a conclusion that is blatantly fallacious for it is clearly a definist fallacy. No doubt it is difficult to read as childish a syllogism as this one in entire 'Utilitarianism', except one assertion in the form of the above mentioned minor premise that the 'desirable' and the 'good' are in essence not two different properties and, this Frankena shows, makes no great harm. Moreover, what reason is there to think that expression like, 'the same as' or 'means' or 'identical' should be taken as 'defined as' or 'proved'?

That Mill has undoubtedly logically proved the first principle 'Utility' stems out of another orthodox evaluation of yet another passage in Mill's 'Utilitarianism'. This time Bradley champions the effort. He thinks that when Mill says, 'No reason can be given why the general happiness is desirable except that each person, so far as he believes it to be attainable, desires his own happiness. This, however, being the fact, we have not only *all the proof which the case admits of* but all which it is possible to require that happiness is a good, that such person's happiness is a good to that person and the general happiness, therefore, a good to the aggregate of all persons. Happiness has made out its title as one of the end of conduct, and consequently one of the criteria of morality',[6] 'our great modern logician thought that by this he had proved that happiness of all was desirable for each'[7] and 'Either Mill meant to argue, because everybody desired his own pleasure, everybody desires his own pleasure or because everybody desires his own pleasure, everybody desires the pleasure of everybody else.'[8]

Bradley's confusion stems out of the expression 'all the proof that the case admits of' and he attacks 'great modern logician' as decisively presenting a logical proof in favour of the principle of 'Utility'. Here 'logical proof' carries a sense of strict logical deduction. Such a confusion again lies in the improper reading of Mill's text. Let us consider the very first sentence of Chapter IV of Mill's book whence he disavows of any strict proof of the principle of Utility. He says, 'It has already been remarked that question of ultimate ends does not admit of proof, in the ordinary acceptance of the term'.[9] Mill clearly rejects any attempt to give a strict logical proof of the principle of utility, but he does favour some other way to establish the first principle (this I shall consider later on).

Moore and Bradley should have noticed the very illuminating title of Chapter IV of 'Utilitarianism', which reads: 'Of what sort of proof the Principle of Utility is susceptible'. It is crystal clear to even one who knows less English than Moore and Bradley that Mill was not giving a proof of the principle of Utility; rather he was more interested in discussing the questions of the meaning of 'proof' when applied to an ethical first principle. He asks 'What ought to be required of this doctrine—what condition is it that the doctrine should fulfil to make good its claim to be believed?'[10] Clearly, Mill carefully overlooks any possibility of strict logical proof of the principle of Utility. Thus, the imputation of provability thesis in the passages quoted by Moore and Bradley in their books is completely unfounded. This basic confusion of Moore and Bradley resulted in a series of confusion. The first is that Mill logically defined 'good' and committed naturalistic fallacy. The second, Mill derived 'ought' from 'is'. Those who have residual doubts about this latter fact should be satisfied now. One should know that Mill was philosophically opposed to any form of a priorism and metaphysicalism in ethics and preferred empiricism. But an empiricist cannot hold that we directly perceive ethical attributes of particular actions. Hence, the establishment of ethical first principle by inductive generalization is ruled out. The neo-evaluators (unlike orthodox evaluators) adhere to an interpretation according to which an empiricist like Mill cannot strictly prove ethical first principle. He cannot prove by deduction, from any more ultimate principle, that there are no unobserved entities, that there are no visible things seen, audible occurrences never heard and so on. But he can set it up as a justified principle that any epistemological theory that requires visible or audible entities that are never seen or heard is a sheer nonsense. The only justification one can seriously propose that a thing is visible is that it is actually seen. By parity of reasoning, an ethical theory that speaks of an end desirable in itself begs that somebody *desires* it in fact otherwise it is just academic and unrealistic. Hence, a psychological realist like Mill demanded that a first principle is incapable of proof and as such the question of *deriving* logically 'ought' from 'is' is ruled out. But unless some justification of the first principle on the basis of psychological realism is set up, the first principle makes no sense. And what could be the basis for such justification? Mill had one option to show that people do *actually desire* as end many things besides pleasure but he makes it clear that despite this 'there is in reality nothing desired except happiness'. Mill is thus successful in showing that it is nothing but utility which is the end principle in question and is acceptable by all, although all *appear* to desire many other ends. Mill wants to show that *only* pleasure is desired for its *own sake* and other things that we seek are parts of happiness. If this is the case, what does it really mean when Mill says that the principle of utility successfully passes the test of psychological realism? It means that happiness is the sole good *not* because pleasantness is good but things

experienced are pleasant, and they alone are good. Pleasure, as a property, is not good, and certainly not the sole good. I think I have reached the brink of understanding the fact that it is needless orthodox thinking to accuse Mill of either providing a strict logical definition of 'good' or a strict logical proof of 'Utility'. Thus, with regard to the general question of proving in ethics, Mill says that no first principle in ethics is in principle provable. As such, utilitarianism as an ethical first principle is *non-provable.* This is is the crux of Mills's non-provability thesis and is central in his ethics. Mill elegantly clarifies a truth and it is that a first principle in an ethical system obviously *cannot* be proved in that system because that begs yet another general principle within that system, which is ex-hypothesis and ruled out. It is due to this that Mill asks a very cogent meta-ethical question "Of what sort of proof is the Principle of Utility susceptible?" His questions were to show the tenability of proving in an ethical system, not that an ethical principle be proved in such and such way. But this does not mean that Mill, like intuitionists Moore and Sidgwick, took recourse to intuitive justification of the first principle in morality. This is evident from his theory of justification, which I shall consider later on.

Once this central theme of non-provability is clear, we are ready to dispel some doubts about Mill's ethics. First is an on-going 'act versus rule controversy', i.e., whether or not Mill is an act utilitarian. I would like to argue that one who fails to understand the non-provability thesis of Mill thinks that Mill is an act and not a rule utilitarian. Again, the misconception troubles those who fall in the line of orthodox thinking. We should be thankful to the epoch-making neo-evaluation of Mill's thesis by J.O. Urmson (*Philosophical Quarterly*, Vol. III, 1953) that has enabled us to dispel such misconception. J.O. Urmson has laid down a criterion for deciding whether or not Mill is a rule utilitarian and has analyzed to show that he is a rule utilitarian beyond doubt. R.B. Brandt (*Ethical Theory*, Englewood Cliffs, N.J. Prentice Hall, Inc., 1959) and W.K. Frankena (*Ethics*, Prentice Hall of India, New Delhi, 1989) adhere to Urmson's criterion.

(i) Brandt criterion: 'Rule Utilitarianism is a method for determining what acts are right.'[11]

(ii) Urmson criterion: 'A particular action is justified as being right by showing that it is in accord with some moral rule. It is shown to be wrong by showing that it transgresses some moral rule'[12] and also, 'A moral rule is shown to be correct by showing that the recognition of that rule promotes the ultimate end.'[13]

(iii) Frankena criterion: 'Rule Utilitarianism emphasizes the centrality of rules in morality and insists that we are generally, if not always, to tell what to do in a particular situation by appeal to a rule. . .'[14]

If we carefully go through all these criteria, we find that Brandt and

Urmson criteria are not any different. They both speak about the fact that rule conformity is in itself a right making characteristic. Frankena criterion needs no separate treatment either. Thus, rule conformity has to be kept in mind in the situation in which an action ought to be done and to choose the most beneficial rule or action in that situation may lead to a beneficial consequence, which is nothing but promoting maximum general welfare. Rule utilitarianism is thus a method for determining what acts are right in a particular situation and what acts may lead to promotion of the maximum general welfare. This *Brandt-Urmson-Frankena* criterion thus tells us that Rule Utilitarianism is a theory according to which "an end justifies moral rules, and moral rules justify acts". On the other hand, act utilitarianism is a moral theory according to which the morality of an action is judged by an act and its total consequences in a particular situation. Following R.B. Brandt again (e.g. "Ethical Theory" 1959), we may put it schematically as:

$\Sigma WP_1 > \Sigma WP_2$, where ΣWP_1 is total welfare consequences in one situation and WP_2, those in another situation. Hence, the total welfare consequence in one situation, if it is greater than those in another situation, it is more desirable or is good, in this sense, act utilitarianism can be aptly paraphrased as a moral theory according to which 'act justifies an end.'

I have in mind two senses of rule utilitarianism, given the above criterion. One is what I call the *primary* or the stronger sense of Rule Utilitarianism, i.e., 'end justifies an act' and the other, the *secondary* or the weaker sense, i.e., 'end justifies a moral rule and a moral rule justifies an act.' Brandt, Urmson and Frankena mostly stick to the weaker sense and are thus troubled by one doubt; whereas the stronger sense, which is implicit in the above mentioned criterion, helps us in dispelling such doubt. Urmson cites the 'exception instances' in Mill's 'Utilitarianism' and says that given these instances, Mill is in apparent contradiction and his rule utilitarian thesis is jeopardized. These instances are (a) cases of *conflict* between secondary principles and (b) cases of the bad consequences of following a secondary rule or of breaking such rule to be good. 'We must remember', writes Mill, 'that only in the cases of conflict between secondary principles is it requisite that the first principle should be appealed to.'[15] And again about the chief exception instance he writes, 'In the case of abstinences indeed of things which people forbear to do from moral considerations, though the consequences in the particular cases might be beneficial—it would be unworthy of an intelligent agent not to be consciously aware that the action is of a kind which, if practised generally, would be generally injurious, and that this is the ground of the obligation to abstain from it.'[16] The problem is if the end principle is what justifies the moral rules and that the principle itself is well established, how can there be conflict between moral rules or secondary principles of morality? And how can a moral rule lead to a bad consequence or that breaking it is desirable? This is a riddle that has

misled many neo-evaluators to think that Mill is in fact an act utilitarian and not a rule utilitarian. But the Brandt-Urmson-Frankena criterion suggests only the fact that rule utilitarianism is tantamount to 'the end justifies moral rules and moral rules justify an act', which is only partially true because Mill has clearly stated that the moral rules are secondary in status in moral judgement. He calls them 'direction posts' or 'sign posts.'[17] We are also thankful to Urmson for helping all neo-evaluators in understanding that the 'exception instances' are deliberately cited by Mill to make it clear that moral rules are not all important in moral judgement, rather it is the first principle or the end principle, i.e., Utility, which is of chief importance in judging the morality of the case. This is the stronger sense in which Mill is a rule utilitarian. Hence, Mill is a rule utilitarian also in so far as he says, 'the end justifies the act' notwithstanding cases of conflict among moral rules, which are only secondarily important in morality. Hence, moral rules are neither necessary nor sufficient conditions for passing moral judgements. A neo-evaluator has to thus understand the fact that Mill is a rule utilitarian in both the senses. However, primarily, he is a rule utilitarian because he never forgets the chief importance of the basic moral principle in moral judgement. Hence, 'exception instances' cited by Mill are not contradictory to his own thesis, rather are used deliberately by him to make clear that after all, 'end justifies act' is the main sense of his rule utilitarian thesis.

This takes us to realize that the central thesis of Mill's ethics, is the non-provability of the ethical first principle; because the principle of utility, which is used to justify everything else (the acts and the moral rules, as well), is itself non-provable and in this sense (primary or strong) 'end justifies act' is what Mill wants to convey in his ethics. Mill says, 'But to consider the rules of morality as non-provable is one thing; and endeavour to test each individual action directly by the first principle is another. It is a strange notion that the acknowledgement of a first principle is inconsistent with the admission of secondary one.'[18]

I think Mill could not have been more clear in stating that the moral rules or the secondary principles are not primary in judging the morality of an action. Hence, the first principle, i.e., 'utility' is what justifies act and in this sense, we call Mill a rule utilitarian. If this be the case, then whether or not Mill is rule utilitarian, gathers no momentum because he never justifies the end by acts and as such, can never be called an act utilitarian. Thus, the 'eclipsing' of the moral rules in moral judgement and placing the first principle to perform the job draws our attention towards the fact that Mill is a rule utilitarian in the stronger sense—'end justifies act.' This truth can be conceived easily if and only if the central thesis of Mill's ethics, that is, the non-provability of the ethical first principle, is given proper importance.

Misled by Urmson's analysis of 'exception instances', a so called neo-evaluator, J.D. Mabbott ('Interpretation of Mill's utilitarianism' in J.B. Schneewind (ed.), *Mill*, Macmillan, London, 1969) said that if moral rules are said to be 'logical corollaries' (p. 25) of ethical first principle, then the 'exception instances' imply that (a) the first principle is *prima facie* duty itself or (b) the basis of such duty or (c) some sort of arbiter between conflicting duties. Mabbott says, 'But when two rules conflict, what question do I ask? And do I apply the first principle to escape my dilemma? Do I ask whether keeping the one rule would in general do more good than keeping the other? This would seem, on Urmson's interpretation, to be the right question, but it would be very difficult to answer. Or do I ask whether keeping the one rule on this particular occasion will do more good than keeping the other? But then I might as well have left out all reference to the rules and just asked whether act A which happens to accord with rule X will do more good than act B which happens to accord with rule Y. Mill gives no guidance to the question which he would approve.[19] He then adds, 'One way of putting the two present difficulties is that on Urmson's interpretation of Mill, the production of the greatest happiness would have to be (a) a *prima facie* obligation; (b) the basis of every other *prima facie* obligation; (c) the arbiter between conflicting *prima facie* obligation.'[20]

But interestingly, Mabbott himself gives us an example in his paper that came very close to recognizing the truth that conflict among moral rules is but natural and that does not justify that the basic principle itself is self-contradictory. He speaks of an honest secretary of a rich man who fell in a dilemmatic situation (conflict) to either take some money from his employer's pocket and give it to a poor man or refrain from it. Mabbott says, in situations like this, one ought to remember the basic moral principle that it may not be useful for maximum people and thus 'you must not weaken your good habits.'[21] He also helps us to realize that Mill's utility principle is just not quantitative pleasure theory, rather a qualitative theory. But Mabbott failed to recognize the secondary status of moral rules and the basic importance of the first principle in morality. Hence, Mill is a consistent rule utilitarian because he never gave up the basic position that the end principle is in essence 'non-provable'. In the strong sense, it is central in judging the morality of an action and as such cannot be said to be a mere arbiter between conflicting duties. However, in so far as Mabbott thinks that the first principle is the basis of every *prima facie* obligation for it justifies secondary rules, he is right. He is also right in saying that first principle as such is a *prima facie* obligation or the basis of the moral rightness of particular acts. Mabbott is very close to understanding the primary or the stronger sense of rule utilitarianism. But he forgets that this does not imply that first principle needs to act as arbiter between conflicting moral rules because in cases like this, we simply refer to the first principle of moral judgement and not to smother any conflict. Thus, Mabbott misses (as do many orthodox evaluators) the central point in Mill's ethics, that is his non-provability theory and the

primary sense in which he is a rule utilitarian. But, Mill's contention that moral rules are the 'logical corollaries'[22] of the ethical first principle raises some doubts. Mabbott initiates this squabble to show that Mill's contention invites inconsistency. Urmson suggests that 'logical corollaries' need not be taken as logical implication but as *heuristic* in nature and this is supported by Mill when he relegates the status of moral rules to 'direction post'. Urmson writes, 'intermediate generalizations' from supreme principle, or as 'corollaries' of it. These are probably the sort of phrases which lead people to think that they play a purely heuristic role in ethical thinking for Mill. . . . It really seems unnecessary to say much of the expression 'corollary'; Mill obviously cannot wish it to be taken literally.'[23] Secondary principles are thus, according to Mill, *heuristic devices* generally justified by the first principle, but not always necessarily beneficial and if they are not qualitatively beneficial for all (as in the case of Mabbott's honest secretary), then we have but one way left for us to decide what to do and that is to refer to the ultimate moral principle, that is, 'maximization of general welfare' where 'general' is *not all* and 'welfare' is *qualitative welfare.* Thus, we reach the conclusion that one who keeps in mind the non-provability thesis of Mill, understands the fact easily that he never justified the end principle with the help of act and its consequences, strictly logically, rather he always spoke of the end justifying an act. Hence, he is a rule utilitarian beyond doubt.

But the last point of our discussion runs into another squabble. If Mill speaks of the *qualitative* end principle or so to say speaks of the role of quality of happiness in the evaluation of the consequences of the action, is it not that he transcends the basic tone of hedonism itself? If this is the case, then certainly Mill is an inconsistent hedonist, as some neo-evaluators like W.K. Frankena believe. Frankena writes, 'John Stuart Mill, partly in reaction, sought to introduce quality as well as quantity into the evaluation of pleasures; but, if one does this, it is hard to see how the utilitarian standard is to be stated, and Mill never makes this clear.'[24] He further adds that 'critics of hedonism often say that Mill's denying it (quantity of pleasure) is inconsistent with his being a hedonist. . . where Mill gets into difficulty is in trying to formulate the principle of Utility in non-quantitative terms.'[25] I would like to justify further that such a misconception again stems out of an improper understanding of Mill's non-provability thesis.

No doubt, Mill would have been obviously inconsistent if he had argued that always the less pleasant of the two pleasures was better or that only the meta-physical pleasures are the better ones. Certainly, he would then be called an inconsistent hedonist. He said, 'it is quite compatible with the principle of utility to recognize the fact that some kind of pleasures are more desirable and more valuable than others. [26] The first thing to notice is that Mill's distinction of the higher and lower pleasures were conceived by the orthodox evaluators (and Frankena in this regard

fell back to such evaluation) as a difference in degree and they overlooked that any difference in degree is also a difference in *kind.* Mill writes, 'The pleasure derived from the higher faculties is to be preferable in kind apart from the question of intensity, to those of which the animal nature is susceptible.'[27] Mill's view is more plausible as the subjects of comparison are not particular pleasures but but a *way of living* or in his words 'modes of existence'. (p. 12) Mill is most convincing when he argues that the existence is more profound in kind than that to be derived from some other modes of existence. This shift in Mill's stance from his predecessors makes the orthodox interpreters like Bradley to remark that 'Mill leaves hedonism altogether.'[28] I must say that Mill no doubt leaves sensationalistic hedonism, but not hedonism as such. Mill's 'theory of life', while incomparably superior to sensationalistic hedonism, is by no means reducible to, say, formalism. Mill, no doubt, lays stress on quality of pleasure but this 'pleasure' is not pleasurable sensation and the 'pain' is not painful sensation. If this is indeed Mill's view, it is far superior to sensationalistic hedonism, without being reducible to formalism. Therefore, I find no sense in thinking that he is an inconsistent hedonist.

However, Mill is of the opinion that in so far as human nature is concerned, he is *capable* of *desiring the desirable,* apart from desiring the desired. Hence, there is qualitative difference between the 'desired desires' and the 'desirable desires.' He says, 'If human nature is so constituted as to desire nothing which is not either a part of happiness or a means of happiness we can have no other proof, and we require no other, that these are the only things desirable.'[29] This reveals the truth that the maximization of pleasures of any kind is not the only end of man, rather he has the disposition to maximize qualitative happiness which may not be desired by all and may not be thus quantitatively beneficial. What is desirable ought to be of utmost *use* to the entire humanity and as such ought to be summum bonum of human life. But is it true that in propagating such an idea, Mill is nourishing formalism? People like Sidgwick, Bradley, Moore and even Frankena suffer from such an idea and I want to reach to the cause of such an idea. The argument of these orthodox thinkers is peculiar. It may be rephrased as, 'since all forms of utilitarianism are forms of hedonism, and rule utilitarianism propagated by Mill is a form of utilitarianism and as such must be hedonistic, which it is not, for it speaks of the quality of pleasures and any consistent hedonistic doctrine rules out ex-hypothesi the concept of quality of pleasures. It speaks of quanitity only.' Given this argument, one wonders why hedonism should mean an ethical theory that recognizes only the importance of maximizing any kind and any quality of pleasures whatsoever to all. If we delve deeper, we find that according to them, act utilitarianism might be called consistent hedonism because it has a distinct hedonistic overtone. Let us consider act utilitarianism first. Even if we accept that act utilitarianism is a consistent form of hedonism, we

are not consistent enough because we ought to qualify our assertion by stating that as a form of hedonism, act utilitarianism is an ethical theory that in principle recognizes the importance of 'pleasure for many', for it is a form of utilitarianism no doubt and must accept this basic truth. If we accept this, we have to accept also that hedonism meant altruistic hedonism only and if we do not accept this, we have to accept that hedonism is just egoistic pleasure theory and then, what hedonism amounts to is difficult to ascertain. The only way is to accept that hedonism is an ethical theory that speaks of the maximization of pleasure no doubt, but not for one, but for maximum. Then, can we really accept the view that the more egoistically pleasure-centred we are, the more we propagate that this is what is ultimately desirable? Are we then, supporters of ethical hedonism? We have long since shunned such a line of thinking and treated Epicureanism as an inconsistent form of hedonism. Interestingly, we had, on the other hand, accepted the truth that utilitarianism is a form of hedonism in the sense that it is basically an ethical theory that speaks of the maximization of pleasures as the summum bonum of life. However, it had been further classified as act, rule and ideal utilitarianism keeping in mind whether 'end justifies act' or 'act justifies an end' and whether or not quantity of pleasure is the only criterion in evaluating the consequences of an action. If this is the case, then rule utilitarianism is a consistent form of hedonism, is the conclusion.

What interests me most is that confusion about Mill being a consistent hedonist stems out of the basic misconception widespread among the orthodox thinkers that Mill had logically proved the first principle or utility (Frankena, however, cannot be accused of this). This is so because these thinkers argued that Mill takes the help of strict inductive proof in order to establish that the ultimately desirable end is utility. As if Mill was arguing in the fashion that, because A,B,C and others desire pleasures, therefore, 'pleasure is desirable for all'. This we have already seen when Bradley retorts (cf. *Ethical Theory*, p. 114 n), against Mill's alleged blatant logical fallacy. One should understand that an expert of inductive logic that Mill really is cannot commit a blatant error like this. How can the generalization 'pleasure is desirable for all' be induced from 'A, B, C and others desire pleasure'? In fact, the misconception lies in thinking that 'pleasure is desirable for all' or 'the desirable end, that is, pleasure is an end for all' is the same as 'all desire pleasure' or 'all are such that they desire the desirable". How can one be so sure of what *all* are doing or may do when Mill says that 'pleasure is desirable for all?' It is a moral statement in which the sense of ought is implicitly present. If he wants to say that A, B, C and others really desire the desirable, i.e., happiness or pleasure (and this we cannot overlook if we are empiricists), we have enough evidence to say that this desirable end ought to be pursued by all. I have in mind the idea that Mill's generalization is not an inductive generalization or a generalization that is an outcome of strict induction.

Rather, it is a moral assertion (in the sense of oughtness) that is essentially *prescriptive* in nature. 'Pleasure is desirable for all' may at best be a moral hypothesis that suggests an end principle for us to be followed in life. However, this may not be viewed as an inductive generalization that is declarative and also prophetic and the basis of which is what A, B, and C really desire. In this, one can see the hopelessness of accusing Mill of committing the fallacy of desiring 'ought' from 'is'. In fact, Mill suggests and *prescribes a truth* that he believes to be beneficial or useful if accepted and pursued. Hence, both the confusions of proving 'ought' from 'is' and proving 'desirable' from 'desired' stem out of one and the same basic confusion that Mill, somehow or the other, proved that such and such is 'desirable' and is true. I, therefore, conclude that the misery of orthodox confusion is the result of their obsession that Mill necessarily advances logical proofs in order to establish the utilitarian end principle as the first principle in morality. It is interesting to notice that although Frankena is strongly opposed to the orthodox view that imputes naturalistic fallacy and definist theory of justification to Mill's ethics; he, however, does not think that Mill is a consistent hedonist in so far as Mill spoke of the quality of pleasure in the evaluation of the consequences of action. To understand the reasons behind this, we should know that Frankena supports a theory of *beneficience* rather than qualitative utilitarianism because qualitative or quantitative utilitarianism as such is an ethical principle which is in fact a *double* principle because on the one hand it tells us to maximize greatest possible balance of good over evil and, on the other hand, tells us to *distribute* this as widely as possible. Hence, the principle of utility rolled into one becomes a principle of beneficience (in so far as it speaks of maximization of greatest possible balance of good and evil) and a principle of justice (in so far as it speaks of just distribution of good). If this be the case, the principle of utility begs a more general principle, the principle of beneficience and as such cannot be an ethical first principle. And as it also begs a principle of justice for its complete explanation, it is to give up pure utilitarianism. Frankena says that, 'In fact, the principle of utility represents a compromise with the ideal, the ideal is to do only good and not to do any harm. If this is so, then the principle of utility presupposes a more basic principle—that of producing good as such and preventing evil. I shall call this prior principle of beneficience.'[30] Frankena appears to come to the conclusion that Mill's inconsistency lies not in speaking about the quality of pleasure as an ethical and principle, rather in the fact that such a theory itself cannot be a first principle at all. The principle of beneficience is primary in obligation. This principle, says Frankena, consists of four things:

(1) One ought not to inflict evil or harm (what is bad);
(2) One ought to prevent evil or harm;
(3) One ought to remove evil;
(4) One ought to do or promote good.[31]

However, he adds to it the principle of justice because he thinks that all our *prima facie* duties cannot be derived from the principle of beneficience for it does not tell us how good and evil are to be distributed. Hence, principle of justice, along with the principle of beneficience should constitute a peculiar *pluralistic* principle in morality. So Frankena advocates a moral theory, essentially pluralistic in nature, which rolled into one, is the same thing as the principle of utility, because he had earlier said that utilitarianism in essence is a *double faced ethical principle* in so far as the principle of beneficience and the principle of justice is coagulated in it. But he charges that notwithstanding the importance of quality of the end principle, utilitarianism itself cannot be the first principle because there is no first principle in morality. However, I do not understand that if the principle of utility is one such ethical principle that essentially speaks of both these truths, why is it that it should be viewed to be begging either of the two or both the principles conceived separately. It is not difficult to realize that all the four features of the principle of beneficience mentioned above and the basic feature of the principle of justice, viz., equal distribution of good are what constitute the principle of utility. The principle of utility is basically an egalitarian principle in morality that makes any sense if and only if 'happiness of the maximum' or 'maximization of general welfare' or 'greatest balance of good over evil' is made correlative to it. The term 'utility' hardly distinguishes 'greatest balance of good over evil' and 'equal distribution of good produced over evil.' Mill writes, '*Social and distributive justice*' is involved in the very meaning of utility, or the greatest happiness principle. The principle is a mere form of words—under one person's happiness, supposed equal in degree—is counted for exactly as much as another's. Those conditions being supplied, Bentham's doctrine, 'everybody to count for one, nobody for more than one", might be written under the principle of utility as an explanatory commentary.'[32] I think that Frankena's alleged attack against Mill that he is an inconsistent hedonist for he speaks of the quality of pleasures, is mainly due to the misunderstanding that a qualitative principle of utility although being consistent enough in so far as it is also a principle of beneficience is inconsistent in so far as it does not go alongside the principle of justice because an egalitarian principle should be such that it need not consider the quality of good produced over evil; it ought to consider the quantity as well when the question of distribution is central. Frankena forgets that Mill never misses the point that after all quantity and quality of good produced over evil is important but he qualifies it by saying that quantity of good produced over evil need not be the same as *any amount* of sensualistic good produced over evil and there is no point in his way of life to distribute any kind of good produced over evil to all—Socrates and fools.

I fear that one who still has residual doubt about Mill's non-provability

thesis may argue that Mill stressed on the fact that virtues are to be imbibed in one's character because they function as means toward attaining the end, i.e. utility. If this was so, he did not support the view that morality primarily consists in framing rules of conduct, notwithstanding the fact that the basic principle in fact might either justify those rules of conduct to be of some value in judging the morality of an action. Thus, an ethics of principle holds that morality consists in framing *rules of conduct* so that there are morally good actions and these actions conform to the rules and then judged 'good' or 'right'; whereas, an ethics of virtue holds that morality consists not in framing rules of conduct, rather cultivating virtues or traits of character. To put it simply, an ethics of virtue propagates the idea 'the act justifies the end'; whereas the ethics of principle propagates the idea, 'the end justifies the act'. Now that we have discussed in detail the role of moral rules in Mill's ethics and the primacy of the first principle in morality, we need only say, with some risk of repetition, that Mill supported the ethics of principle. However, Mill also believes that morality does not merely consist in framing rules of conduct, rather *following* it as well. Otherwise, morality is barren; our disposition to follow these rules and thus 'desire the desirable' is crucial. This is evident when Mill says, 'a necessary part of a perfectly just conception of utility or happiness, is considered as the directive role of human conduct.'[33] He makes clear that utility as a first principle in morality is of great value if and only if one follows it not for his own sake but for others as well because a noble character is one who makes efforts to make other people happier. Hence, one who keeps the basic moral principle in mind that it justifies all moral actions, never forgets that 'Utilitarianism, therefore, could only attain its end by the general cultivation of nobleness of character.'[34] Hence, the cultivation of the nobleness of character is important in a utilitarian scheme, but what is primary is that such nobleness of character makes sense only if the basic moral truth is kept in mind. One who still gets stuck to the idea that, for Mill, capacity to follow certain rules is what is crucial, and, therefore, imbibing virtue is primary, forgets that following moral rules makes *no sense if the ethical first principle* is not recognized as a moral standard. It is this principle that justifies moral rules and not the other way round. Hence, a supporter of the ethics of virtue is stuck to the idea that Mill proves that such and such moral rules ought to be followed as these rules are intrinsically valuable and also desirable. Moreover, it is also to be believed that following moral rules is what is the foundation of Mill's utilitarian thesis. However, moral rules are mere 'direction posts' that make any sense if and only if they are directed to achieve an ultimate end in morality. This end is intrinsically valuable and desirable and as moral rules derive value from the end principle (as heuristic devices), they are also important but we may dispense with these rules (in exceptional cases) to refer to the end principle in question. Mill's ethics, thus,

irresistibly preaches the importance of but one thing and that is, the ultimate moral principle, viz., utility. Hence, his is an ethics of principle.

From the overall discussion, I want to draw three broad conclusions: one is that Mill never advanced any logical proof to establish that the first principle in morality is utility; secondly, this non-provability thesis is central in Mill's ethics in understanding that he advocated rule utilitarianism, which is consistent hedonism, and finally, that Mill's ethics is an ethics of principle and not of virtue. But there is a problem still haunting all of us who want to get a clear answer to the important meta-ethical question raised by Mill in Chapter IV. After all, what sort of proof is the ethical first principle susceptible? Mill must face the charge of intuitionism if he has no pertinent answer to this problem. It is here that Mill advances the notion of *informal justification* in ethics. It is according to me a theory of 'psychological justification' that speaks of some non-logical sense in which our basic norms and value judgements can be justified by appeal to the nature of things. But 'non-logical' does not mean illogical; rather, not involving strict deductive or inductive logical rules. A theory may be heuristically developed. We also speak of the role of informal logic in ethical justification. Hence strict logic may not be the only way to establish that such and such is a justified moral theory. This scope of establishing a moral theory in a non-logical way takes us to Mill's theory of 'psychological justification'. The theory is based on two important presuppositions: (a) no strict logic can prove an ethical first principle; (b) any justification of an ethical first principle starts from recognizing the fact that human nature is so constituted that ultimate end of life is nothing but maximizing the general welfare; for one who is pressed upon to give reasons for the claim that human nature is constituted in the way mentioned above, the answer is that the psychological appeal to the *honesty* of a man to *accept* this truth is what is more important than giving one reason or another. Mill's psychological justification has its distinctive mark in the fact that 'appeal to intellectual honesty' is what is most crucial in this regard and I think it is in this sense that his theory of justification is different from that of the definist or the intuitionists and the extreme non-cognitive theories of justification. Mill says, 'We are not, however, to infer that its (the first principle) acceptance or rejection must depend on blind impulse or arbitary choice. There is a larger meaning of the word "proof" in which this question is as amenable to it as any other of the disputed questions of philosophy. The subject is within the cognizance of the rational faculty and neither does that faculty deal with it solely in the way of intuition. Considerations may be presented capable of determining the intellect either to give or withhold its assent.'[35] Clearly, Mill did not prefer to logically prove (either deductively or inductively) that the first principle is utility and, therefore, he did not provide us with any logical theory of justification that is in fact definist in nature. Also it is a non-intuitionist theory of justification in so far as he did not take

recourse to self-evidence and intuition in establishing a moral theory. He also does not propagate a theory of negative justification akin to that of extreme non-cognitivism because he did not believe that any justification whatsoever of a moral principle is at all possible and, therefore, an ethical principle has to be simply accepted to be true. Thus, Mill did not prefer definist theory of justification because he was opposed to providing any kind of definition of the moral concepts in natural terms and in no way derived 'ought' from 'is'. We can easily see that Mill's theory of justification is a result of his non-provability thesis and we can understand why his non-definist stance of justification enables us to say that he was a naturalist who still refrained from giving any naturalistic definition to the moral term 'good'. Interestingly, Mill's naturalism needs careful understanding because 'naturalism' is not an ethical view according to which an ethical term and a natural term are necessarily identified; rather, it is an ethical view according to which, (a) the moral term 'good' *makes any sense* if and only if it is *constantly conjoined* to another 'concept', i.e., Utility, (b) 'Utility' *makes any sense* if and only if 'maximization of general happiness' is constantly conjoined to it, and (c) in moral justification, acts and their consequences stand out to be crucial *evidences* but not logical grounds for deriving the end principle. Acts and their consequences are thus not logical grounds for justification; rather, *evidence* in course of psychological justification in which *self-observation, self-consciousness* and *observation of otherselves* are crucial. Mill says, 'And now decide whether this (mankind desires happiness only) is really so. . . . We have evidently arrived at a question of fact and experience dependent upon evidence. It can only be determined by practised self consciousness and self-observation assisted by observation of others.'[36] We can thus legitimately conclude that Mill's non-provability thesis culminates in a theory of justification that makes clear beyond doubt that he is a rule utilitarian (in both senses), a supporter of ethics of principle and a consistent hedonist.

[It has been argued in this paper that the central theme of Mill's utilitarianism is his non-provability theory, which speaks of the impossibility of strict logical proving of the ethical first principle, viz., 'utility'. One who keeps this theory in mind can easily understand the hopelessness of the interpretations and criticism of Mill's ethics by some orthodox evaluators. The neo-evaluators ought to justify that Mill is a rule utilitarian, a consistent hedonist and that his ethics is an ethics of principle because Mill advocated the non-provability theory and the theory of informal justification in ethics.]

Notes and References

1. J.S. Mill, *Utilitarianism*, H.B. Acton (ed.), Everyman's Library, London, 1992, p. 36.
2. F.H. Bradley *Ethical Studies*, Oxford, 1927, p. 115.
3. G.E. Moore, *Principia Ethica*, Cambridge University Press, London, 1903, p. 67.

4. Ibid., p. 66.
5. W.K. Frankena, 'The naturalistic fallacy', *Mind*, XLVIII (new series), 1939, pp. 464–77.
6. J.S. Mill, op. cit., p. 36.
7. F.H. Bradley, op. cit., p. 113.
8. Ibid., 114–14n.
9. J.S. Mill, op. cit., p. 36.
10. Ibid., p. 36.
11. R.B. Brandt, 'Ethical theory' Englewood Cliffs, 1959, p. 253.
12. J.O. Urmson, 'Interpretation of moral philosophy of Mill', *Mill*, J.S. Schneewind (ed.), Macmillan, London, 1969, p. 183.
13. Ibid., p. 183.
14. W.K. Frankena, *Ethics*, second edition, Prentice Hall of India, New Delhi, 1989, p. 33.
15. J.S. Mill, op. cit., p. 27.
16. Ibid., p. 25.
17. Ibid., p. 20.
18. Ibid., p. 25.
19. J.D. Mabbott, *Mill*, J.B. Schneewind (ed.), op. cit, p. 192.
20. Ibid., p. 193.
21. Ibid., p. 196.
22. J.S. Mill, op. cit, p. 25.
23. J.O. Urmson, op. cit, pp. 186–87.
24. W.K. Frankena, op.cit., p. 36.
25. Ibid., p. 84.
26. J.S. Mill, op. cit., p. 8.
27. Ibid., p. 11.
28. Bradley, F.H., op. cit., p. 115.
29. J.S. Mill, op. cit., p. 40.
30. W.K. Frankena, op. cit, p. 45.
31. Ibid., p. 47.
32. J.S. Mill, op. cit, p. 65.
33. Ibid., pp. 12–14.
34. Ibid., p. 13.
35. Ibid., p. 5.
36. Ibid., p. 40.

Two Kinds of Relative Motion: An Interpretation of Berkeley's Distinction Between T-motion and A-motion

PRAJIT K. BASU
Department of Humanities and Social Sciences,
Indian Institute of Technology, Delhi

It is well known that George Berkeley, the famous eighteenth century British philosopher, held that all motion is relative. Berkeley's attack against absolute space and absolute motion rests on his contention that motion is relative by its very nature. In this paper, I shall first analyze the distinction that Berkeley had attempted to draw, in *A Treatise Concerning the Principles of Human Knowledge* (hereafter *Principles*), and in *De Motu* (hereafter *DM*) between a body in motion or in apparent relative motion (hereafter A-motion) and a body moved or a body in true relative motion (hereafter T-motion). I show that Berkeley's texts give us two methods for verifying whether a body is in T-motion. Next, I explore the role of force as the ground for the distinction between the A-motion and the T-motion of a body. In the next section, I employ this result to interpret Berkeley's analysis of motion in the context of impact phenomena. I also argue that the role of reference frame from which the T- or A- nature of motion of a body is determined, is extremely important. In the last section, I show that K.P. Winkler's interpretation of Berkeleyian analysis of motion fails to handle some cases of impact phenomena adequately.[1] Winkler's suggestion of 'force-endowing' history[2] of a body in T-motion as an adequate analysis of the distinction between T-motion and A-motion, I argue, is to be extended by an analysis which is simpler, compatible with Berkeley's text, and the Principle of Inertia, and is sensitive to the role of reference frame in determining both the magnitude and the direction of motion of a body.

SECTION I: TRUE MOTION AND APPARENT MOTION

Berkeley takes four concepts as basic in understanding physics. These are moved, moving, motion, and rest. This is made clear in section 11 of *De Motu* where Berkeley argues that there is nothing to which dead force or gravitation refer. This section also makes clear how one can understand the concept of force. Berkeley says,

> . . . those terms dead force and gravitation by the aid of metaphysical abstraction are supposed to mean something different from moving, moved, motion and rest, but in point of fact, the supposed difference in meaning amounts to nothing at all.[3]

It must be noticed that Berkeley is making a distinction among moving, moved, motion and rest, although he does not tell us here in what the distinction lies. In fact, no where in *De Motu* does he make the distinction clear. In all fairness, however, he may be taken to accept the distinction between motion and rest of a body which was prevalent at that time. For him, motion and rest are two different states of a body in which a body can persist. Motion is the successive existence of body in different places.[4] His acceptance of the Principle of Inertia as a primary law of nature commits him to accept that a state of rest can be changed to a state of motion and vice versa. There is enough textual evidence to show that he accepts the Principle of Inertia and the distinction between motion and rest. For example, he says,

> . . . it is a primary law of nature that a body persists exactly in a state of motion and rest as long as nothing happens from elsewhere to change that state.[5]

Berkeley does not give any method to distinguish between moving and moved or between moving and motion (or rest) or between moved and motion (or rest). We do get some idea of these distinctions in *Principles.* Berkeley introduces the distinction between motion and moved in three sections of the *Principles.* He says,

> . . . though in every motion it be necessary to conceive more bodies than one, yet it may be one only is moved, namely that on which the force causing the change of distance is impressed, . . . that to which the action is applied.[6]

It appears that Berkeley introduces a distinction between a body in motion (i.e., in A-motion) and a body moved (i.e., in T-motion). A body, X, is in A-motion if and only if the relative position of X with respect to another body (say Y) is changed. The change in relative position is a necessary and sufficient condition for anything to be in motion. X is in T-motion if and only if

(1a) the relative position of X with respect to another body (say Y) is changed
(1b) a force is impressed upon X or an action is applied to X; and
(1c) the force impressed or the action applied causes the change of relative position of X.

The asymmetry between X is in A-motion and X is in T-motion is obvious. The change of relative position of X with respect to Y is both a necessary and sufficient condition for apparent (or A-) motion while it is only a

necessary condition for true motion. That both (1b) and (1c) are necessary conditions for X to be in true (or T-) motion is emphasized by Berkeley by rejecting the thesis that the definition of X is in T-motion is given by only the condition (1a). Thus, Berkeley says,

> For . . . some may define relative motion, so as to term that body moved, which changes its distance from some other body, whether the force or action causing that change were applied to it or no: yet as relative motion is that which is perceived by the sense, and regarded in the ordinary affairs of life, it seems that every man of common sense knows what it is . . . now I ask anyone whether in his sense of motion as he walks along the street, the stones he passes over may be said to move, because they change distance with his feet?[7]

Berkeley's rhetorical question is to only point out that mere relative motion, i.e., change of position, is not a sufficient condition for saying that X is in T- motion. That all three conditions need to be satisfied for a body to be in T-motion is argued by Berkeley in the following quotation:

> So on the other hand, when one only body (the parts whereof preserve a given position between themselves) is imagined to exist; some there are who think that it can be moved in all manner of ways, though without any change of distance or situation to any other bodies; which we should not deny, if they meant only that it might have an impressed force, which upon bare creation of other bodies, *would* produce a motion of certain quantity and determination.[8]

Berkeley's point is that when there is only one body it is not in any kind of motion. It is possible that a force can be applied to that body (that satisfies condition 1b). That force would be the cause of the motion of that body (thereby satisfying condition 1c) provided there is another object with respect to which that motion is to be produced (thereby satisfying condition 1a). Also, since motion is relational, any talk about force acting on a single body must be in terms of certain subjunctive statements like if there were another body, then the body X upon which the force is acting would be in T- motion with such and such magnitude and such and such direction. This statement is in conformity with Newtonian laws (including the Principle of Inertia) and for Berkeley it is justified by observing the regularities obtained between the force applied and the motion produced in actual cases. The above quotation thus serves as a textual evidence for Berkeley's claim that condition (1b) and (1c) are each necessary for a body to be in T-motion.

That there is a crucial distinction between X is in T-motion and X is in A-motion is driven home by an analogy:

> To me it seems, that though motion includes a relation of one thing to another, yet it is not necessary that each term of the relation be

> denominated from it. As a man may think of somewhat which doth not think, so a body may be moved to or from another body, which is itself not in motion.[9]

The point Berkeley makes is straightforward. If motion is relative, one may say that Y is also in motion with respect to X after all it is the change of position of X (or Y) with respect to Y (or X) which is necessary and sufficient for claiming that X (or Y) is in motion. Berkeley says that so long as one claims that X (or Y) is in A-motion one is not making any mistake. But, only one of them is moved, i.e., is in T-motion. If X is in T-motion, then the distinction between the motion of X and that of Y is as fundamental as a thinking being and the object of thought. A sentiment being S indulging in thinking activity thinks of an inanimate object O. The relation between S and O is asymmetrical. Similarly, X upon which a force is impressed changes its position with respect to Y (which may be considered to be at rest) and is in T-motion. The relation between motion of X and of Y is asymmetrical. The burden of the distinction is borne by force or action that is applied to X. That there is an asymmetry in the relation X is in T- motion is clear from the following quotation:

> . . . to denominate a body is moved, it is requisite, first, that it change its distance or situation with regard to some other body: and secondly that the force or action occasioning that change be applied to it. If either of these be wanting, I do not think that agreeably to the sense of mankind, or the propriety of language, a body can be said to be in motion.[10]

Berkeley means by motion in the above quote the true motion of a body. It needs to be pointed out that Berkeley does not mention about force causing the change of motion. Rather the force or action is only the occasion of such a change.

Berkeley also says that one can be mistaken if a body is in T-motion or in A-motion:

> . . . it is possible for us to think a body, which we see change its distance from some other, to be moved, though it have no force applied to it (in which sense there may be apparent motion), but then it is, because the force causing the change of distance, is imagined by us to be applied or impressed on that body thought to move. Which indeed shews we are capable of mistaking a thing to be in motion which is not. . .[11]

Again, Berkeley says that just because we can make mistake regarding the nature of motion of a body, it

> does not prove that, in the common acceptation of motion, a body is moved merely because it changes distance from another; since as soon as we are undeceived, and find that the moving force was not communicated to it, we no longer hold it to be moved.[12]

Berkeley's point seems to be that we can know whether X is in T-motion once we know that a force or action is applied on X. This is important because it is the knowledge of the application of force on X that grounds the distinction between the knowledge that X is in T-motion and the knowledge that X is in A-motion. The distinction is not grounded in the nature of motion. It is a matter of fact if force is applied to X or not. Hence, it is a matter of fact whether X is in true motion or not. And it is the prior knowledge of the force acting on X that helps to decide the knowledge of nature of motion. Therefore, one way to verify whether a body is in T-motion or not is to verify whether all three conditions (1a)—(1c) are satisfied. This is one of the two ways to verify whether a body is in T-motion. I shall discuss the second method later. Thus, in order to verify whether a body is in T-motion, one has to verify whether a force has acted on it or not (condition 1b). Now, the question is how do we understand force? Berkeley seems to think that it is possible for force to act upon a body and yet the body may be unmoved.[13]

Thus far we have found that there is a distinction between a body in T-motion and a body in A-motion which can be understood in terms of force. We have also mentioned that Berkeley does not like to introduce force as a causal agent. At the most force can be a mathematical entity the existence of which does not have to be asserted. How can we understand force then? Berkeley's answer is that force can be understood only in terms of motion. Thus, Berkeley says,

> . . . no force is immediately felt by itself, nor known or measured otherwise than by its effect.[14]

In section 11 of *De Motu* Berkeley is more specific. He discusses the case of force of gravitation and generalizes the result.

> The force of gravitation is not to be separated from momentum; but there is no momentum without velocity; again velocity cannot be understood without motion, and the same holds therefore of the force of gravitation. Then no force makes itself known except through action, and through action it is measured; but we are not able to separate the action of a body from its motion.[15]

There are quite a few questions which can be asked. How can we measure the action of X? Is it through the motion of X? Let us consider impact between two bodies, X and Y. If Y is in T-motion then Y is in motion at least with respect to some body (say X for the sake of simplicity) and X must have applied a force on Y. The force that X applies on Y can be measured only through the action *of* X and not action *on* Y. The force of X on Y can be measured only by the motion of X. The questions then are: Is it the A-

motion of X before impact or A-motion of X after impact? Or is it the change of A-motion of X due to the impact? Which reference frame is one supposed to employ to determine (the change of) any motion of X? Before I consider these questions, I want to discuss another method, implicit in Berkeley's account, to verify whether a body is in T-motion or not. Consider the example that Berkeley gives:

> . . . as long as a heavy body changes the shape of a piece of lead put under it, . . . so long is it moved, but when it is at rest, it does nothing.[16]

The example above, according to Berkeley, is a case of percussion, i.e. of impact. The example appears to reveal a connection between a body's acting and its T-motion. If there is a change in the shape of a piece of lead (call it Y), i.e., if Y is in motion, then the heavy body (call it X) is moved. Also, if X is at rest, X does not act. What does this example really tell us? Is it the case that change in the shape, i.e., the motion in Y is a sufficient condition for claiming that X is moved in the case of impact? It appears that three different claims can be made which reflect the condition for X is in T-motion and X is at rest.

(2) If X is at rest, then X does not act.
(3) X is in T-motion with respect to an inertial frame of reference, *I*, if and only if
  (a) X is in A-motion with respect to *I*
  (b) X acts on Y and
  (c) puts Y in A-motion with respect to *I*.
(4) X is in T-motion with respect to *I*, if and only if
  (a) X is in A-motion with respect to *I*
  (b) X acts on Y and
  (c) puts Y in T-motion with respect to *I*.

The analysis (4) follows from (1) and (3). For Y to be in T-motion, Y must be in A-motion (condition 3c), and Y must have been acted upon (condition 3b). The difference between the analysis (1) on the one hand and the analysis (3) or (4) on the other is that the former is supposed to be a more general analysis and the latter is a specific one applicable to impact phenomena of bodies. However, (1) could also be directly pressed into service in cases involving impact phenomena as will be discussed shortly. It must be noticed that the condition (1b), X acts on Y, was supposed to be explicated by the example in Section 11 of *De Motu*. But the condition has reappeared in the subsequent analyses (3b and 4b). To summarize, we now have two methods given by analyses (1) and (4) for verifying whether a body is in T-motion, both of which employ the notion of force or action which remains to be explicated. In the next section I employ these methods to some specific cases of impact phenomena.

SECTION II: THE VERIFICATION OF T-MOTION OF A BODY IN IMPACT AND THE ROLE OF INERTIAL REFERENCE FRAME

Suppose there are only two objects (X and Y) in the universe. Suppose X is at rest with respect to an inertial reference frame, *Iy*, fixed on Y. Now let a force act on X. Let us say that X changes its position with respect to *Iy*. Thus, X is in A-motion with respect to *Iy*. According to analysis 1, X is in T-motion. However, the reference frame, *I*, could be fixed on yet another object (for example, the fixed stars), Z, provided the universe was populated by only three objects, X, Y, and Z. In that universe, the analysis of X is in T-motion remains unaltered. Now, let us extend the analysis to the phenomena of impact between X and Y in a universe with only two objects X and Y. Suppose the inertial reference frame *Iy* is fixed on Y. Suppose X is in A-motion with respect to *Iy* (and hence with respect to Y) and the motion is directed towards Y. Let X and Y undergo an impact. After the initial period of disturbances let us assume that X is in A-motion with respect to *Iy* (and hence with respect to Y) but the motion is directed away from Y possibly with a different magnitude than before the impact. The question is: Is X in T-motion? The scenario can be redescribed from the vantage point of the inertial reference frame, *Ix*, fixed on X. The question then is: Is Y in T-motion after the impact? The answer to both these questions are affirmative if Newton's first law (i.e., the Principle of Inertia) is accepted and the analysis (1) is applied. In fact, the inertial reference frame plays a major role in raising and answering these questions about the T- or the A-motion of a body. Described from *Ix*, Y is in T-motion but we cannot say anything about the motion of X, while described from *Iy*, X is in T-motion but we cannot say anything about the motion of Y. And described from *I*, reference frame fixed on a third body (for example, fixed stars), both X and Y are in T-motion. The example can be modified such that with respect to *I*, X is in T-motion but not Y or Y is in T-motion but not X and so on. In the case of impact, then, the description of the nature of the motion of a body is dependent upon the vantage point from which it is described. How do the analyses (3) and (4) of T-motion fare in the cases of impact? In the context of impact phenomenon in a universe with two objects and the reference frame, *Iy*, fixed on Y, the condition (3c) or (4c) cannot be satisfied. Hence, whether X is in T-motion cannot be determined but neither can one determine whether Y is in any motion at all. Similarly, if the reference frame, *Ix*, is fixed on X, again it will be impossible to determine whether X or Y is in T-motion. What does Berkeley's analysis of T-motion then amount to? A careful look at the quotation in Section 11 of *De Motu* gives us an answer. The inertial frame from which the situation of the change of the shape of piece of lead is described is neither on the heavy body nor on the piece of lead. The inertial frame, *I*, is on a body distinct from the heavy body and the piece of lead. The last two analyses, (3) and (4), are then to be

employed from the vantage point of *I*. This is because at least in some cases of impact phenomena, the force applied by one body on the other is cashed out in terms of motion of both bodies. One way to describe the motion of both bodies is to describe from the reference frame fixed on a third body. In that case, the given situation of impact between X and Y would lead to the result that X is in T-motion after the impact. This analysis then provides yet another way of verifying whether X is in T-motion, with respect to a reference frame. However, Berkeley could have meant the following analysis of X is in T-motion in the case of impact phenomenon of the type described in Section 11 of *De Motu*:

(5) X is in T-motion with respect to *Iy*, if and only if
  (a) X is in A-motion with respect to *Iy*
  (b) X and Y are in impact
  (c) X puts Y in A-motion with respect to *Ix*

Notice, however, that both (5a) and (5c) are satisfied even before the impact in somewhat different cases where two bodies are approaching each other. In that case, the distinction is grounded by the condition (5b), i.e., whether there was an impact or not. And that does not seem to be compatible with what Berkeley says. However, we can say that after the impact X is in T-motion because X after the impact puts Y in an A-motion relative to *Ix* which is different from the A-motion of Y relative to *Ix* before the impact. The analysis then requires modification of (5c). The new analysis (6) while retaining (5a) and (5b) would be augmented by (6c): X changes the A-motion of Y with respect to *Ix* before and after the impact. This implies first that change in (magnitude and/or direction) is requisite to determine whether a body is in T-motion. But that should be obvious. The Principle of Inertia tells us that the change in motion is due to a force. Thus, the new motion is a result of the force that acted. And that automatically implies T-motion from the analysis. The analysis also implies that we can make sense of X is in T-motion only from a certain inertial reference frame. This in turn implies that Berkeley perhaps held that there may not be after all any major difference between a body X is in A-motion and X is in T-motion for the difference amounts to only a difference in coordinate frame from which the motion of X is described. K.P. Winkler also believes that Berkeley's analysis underscores the role of reference frame. Winkler argues that an object could be truly in motion and at the same time truly at rest, are from Berkeley's view point always relative. 'The apparent conflict is rooted in alternative (but not conflicting) choices of reference frame.'[17] That Berkeley's view is indeed close can be seen from the following textual evidence. Berkeley says that 'the motion of the same body may vary with the diversity of relative place.'[18] He goes on to say that a thing 'can be said in one respect to be in motion and in another respect to be rest'.[19]

The analysis in the context of a reference frame, *I*, fixed on a body separate from X or Y is then as follows:

(7) X is in T-motion with respect to a reference frame, *I*, if and only if
  (a) X is in A-motion with respect to *I*
  (b) X acts on Y (i.e., X is in impact with Y)
  (c) The impact induces change of motion in Y with respect to *I*.

The part 'X acts on Y' in (7b) must be cashed out in terms of motion of X with respect to *I*. There is some minor problem with such a condition. Consider a ball (Y) hanging from a string. Let us say another ball (X) which is in motion relative to Y and *I*, and has been acted upon before, now strikes Y and imparts motion to Y with respect to *I*. Given the above condition, X is in T-motion. Now consider a situation where Y is not on the direction of X's motion. Then, we will have to say X is not in T-motion. To get around this problem one can say that,

(8) X is in T-motion with respect to *I*, if and only if
  (a) X is in A-motion with respect to *I*
  (b) X can act on Y
  (c) The impact thereby can induce change of motion in Y with respect to *I*.

There are then two methods to verify whether a body is in T-motion with respect to a reference frame in cases of impact. These are:

(9) X is in T-motion with respect to a reference frame, *I*, if and only if
  (a) X is in A-motion with respect to *I*
  (b) X has been acted upon (i.e., X was in impact)
  (c) The impact is the occasion for X's Ā-motion (or change of Ā-motion) with respect to *I*.

(10) X is in T-motion with respect to a reference frame, *I*, if and only if
  (a) X is in A-motion with respect to *I*
  (b) X can be in impact with Y
  (c) The impact would be the occasion for Y's A-motion (or change of A-motion) with respect to *I*.

Several features stand out in this analysis. First, the problem of verifying or knowing whether a body is in T-motion or in A-motion is solved by one of two ways in the case of impact phenomenon. The first way is given by the analysis (9). In this analysis, one observes an impact phenomenon or collision between two bodies and observes the subsequent change in magnitude and/or direction of motion of bodies after the impact. If the body under investigation does undergo such a change, it

is in T-motion after the impact provided it is in A-motion after the impact. The second way would require observing whether the body under observation (X) changes the magnitude and/or direction of motion of the body (Y) with which X collided. If the collision led to such a change in Y then X was in T-motion before the impact (provided X was in A-motion before the impact) and if X is in A-motion now, it is still in T-motion now. Both these methods are compatible with the Principle of Inertia.

Second, the reference frame with respect to which motion is determined is very important especially for the purpose of verifying whether a body is in T-motion or in A-motion. Third, the analysis shows that Berkeleyian analysis fails to bring out the difference between the natures of T- and A-motion in terms of force because it fails to give an account of the nature of what a force is. Fourth, the analysis helps us determine whether bodies in impact were in T-motion before the impact at least in some cases.

SECTION III: APPLICATION OF THE ANALYSIS IN IMPACT PHENOMENA

I will now apply the analysis in certain test cases and draw some important conclusions from the results of these applications. K.P. Winkler in his analysis of Berkeley's theory of motion makes two assumptions. First, when two bodies have different velocities with respect to a reference frame and subsequently undergo impact at least one of them was in T-motion with respect to that reference frame before the impact.[20] Second, in cases of impact Berkeley is committed to hold that whatever applies a force must be in T-motion. Finally, Winkler argues that to say that a body, X, was acted upon by a force, is to say that X has a force-endowing history. In other words, Winkler's analysis requires that the distinction between a body in T-motion and a body in A-motion is grounded by whether there was an event in the past history of the body (now in T-motion) involving it—a collision for example, in the case of impact phenomenon.[21] Once we know that there was such a phenomenon, we know that the body was or is in T-motion. I will argue now that in spite of some advantages Winkler's arguments and assumptions are problematic.

Winkler raises a problem in verification of a body being moved in the case of impact. Assume a body, D, is in motion with respect to an inertial frame, A. The question is how can we know whether D is in T-motion. Given the Berkeleyian distinction between a body in A-motion and a body in T-motion, D is in T-motion if and only if D is struck by a force exerted by a body B. Winkler assumes that B must also be in T-motion to be able to exert a force on D, i.e., B's being in T-motion is a necessary condition for B's being able to act. Now, how can we know that B is in T-motion? B is in T-motion if and only if B is struck by a force exerted by another body C that is in T-motion. Winkler now argues that one has no way to determine if A is in T-motion because one can never have knowledge of

the first body that is in T-motion. Winkler now solves the problem by showing that either B or C must be in T-motion prior to B's striking A and also prior to C's striking B. Let me quote Winkler in detail:

> We are inclined to say that the change in relative position is due to the collision of D with B. But in order to know that B impressed a force on D we need to know that A was in true motion. We know that an even earlier point B and C collided and that before and after the collision both were changing position relative to I. But we cannot say, of either B or C, that it was in true motion before the collision. Suppose for the sake of the argument that neither was. This supposition is impossible because their changes in position relative to I were different; both changes could not have been the artifacts of the motion of I. It follows that one of them must have been in true motion, though we have no way of telling which it was. Fortunately, we do not need to tell: if either was in true motion before the collision then both were in true motion after the collision, because the body in true motion would impress a force on the other which would in turn impress a force on it.[22]

Winkler's claim is that since B and C have different relative motion with respect to A, one of them must be in T-motion. This is the crucial premise for which Winkler's argument is not clear. He does not say why B and C cannot have different relative velocities with respect to I and still not be in T-motion. Does it follow from the nature of motion? Winkler does not explain how if at all it does. I shall return to this point. Winkler's argument then may be summarized as follows. In impact, the impacting bodies must have different motion with respect to a reference frame. At least one of them is in T-motion (first assumption). After the impact, the motion of the bodies involved change and they are in T-motion. The collision, therefore, ensures, by appealing to Newton's third Law, that these bodies are in T-motion. However, if it is the collision which ensures that the bodies are in T-motion after the impact, then we do not need to go back to the impact between B and C to determine whether D is in T-motion after its impact with B. Because the change in motion of D, subsequent to its collision with B, would require a force acting on it as required by the Principle of Inertia. Hence, we do not need to assume that at least one of B or C must be in T-motion. This simplifies the account and is compatible with Winkler's requirement of 'force-endowing' history of a body and does not invoke 'force' as an ontological entity. The analysis (9) or (10) in this sense (of expelling force from the account) is compatible with Winkler's analysis. Moreover, the analysis is compatible with physics to the extent that it holds that impact phenomenon is a reference frame invariant fact of the matter and that the laws of impact give us the rules of exchange of motion among bodies.

However, Winkler's account is problematic for a different reason.

Suppose we assume that the talk about force is cashed out in terms of impact and rules of (ex)change of motion among bodies in impact. Now consider the two diagrams below. Fig (a) shows a snap shot of motions of X and Y at time $t=t_0$ with respect to an inertial reference frame, *I*. Fig (b) shows a snap shot of motions of X and Y with respect to *I* at a later time (after the impact between X and Y) $t=t_1$. In this case, although X and Y have different velocities (in fact Y is at rest), with respect to *I*, and subsequently undergo impact we cannot determine whether X was in T-motion with respect to *I* before the impact. That X is in T-motion, after the impact, with respect *I* is determined by the first method (analysis 9) which follows from the Principle of Inertia. This shows that Winkler's second assumption does not hold for all cases. This is because that X is in T-motion after the impact can be determined from the Principle of Inertia. But we cannot determine that by employing Winkler's analysis. Because for X to be in T-motion after the impact, X must have already been in T-motion before the impact or must have been acted upon by Y which must be in T-motion.

Figure (a)

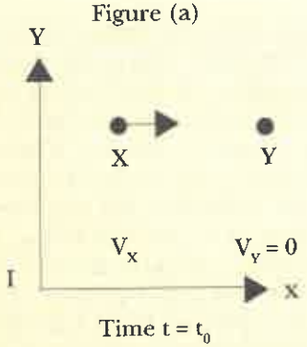
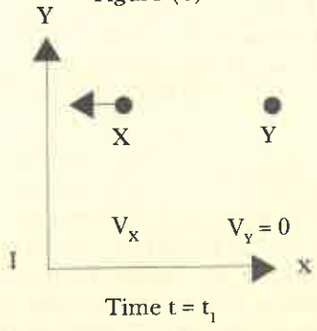


Figure (b)



Is X in T-motion wrt I at t_0?	Undetermined	Was X in T-motion wrt I at t_0?	Undetermined
Is Y in T-motion wrt I at t_0?	No	Ix X in T-motion wrt I at $t=t_1$?	Yes
		Is Y in T-motion wrt I at $t=t_1$?	No

But Y is not in motion. Therefore, X must have been in T-motion. However, we cannot determine that without the analysis becoming circular. Suppose we assume instead that since X and Y have different velocities with respect to *I*, one of them must be in T-motion (first assumption). Since X is in A-motion and Y is at rest, X must be in T-motion. This follows from the first assumption. But as discussed above there is no argument for this assumption. It is an ontological thesis which must follow from the nature of motion and argued for separately. This case of impact thus constitutes a problem for Winkler's assumption. This

is because the Principle of Inertia would commit Winkler to hold that X is in T-motion by virtue of the fact that the change in the direction of X's motion is due to a force acting on it. But his analysis which requires commitment to the second assumption fails to derive the required conclusion unless it commits to a thesis which has not been argued for. My analysis while it can accommodate this case is also in agreement with Winkler's thesis of force-endowing history in the following sense. The analysis acknowledges the fact that in the case of impact, the knowledge of collision would help determine whether a body was or is in T-motion. This follows from the Principle of Inertia. My analysis also helps determine whether bodies are in T-motion in some cases from the data obtained after the impact. In this sense it is an extension of, and improvement on, Winkler's analysis.

Notes and References

1. K.P. Winkler, 'Berkeley, Newton and the Stars', *Studies in History and Philosophy of Science*, 17, 1986, pp. 23–42.
2. Ibid., p. 31.
3. G. Berkeley, *Philosophical Works*, Dent, London, 1983, Sec. 11, *De Motu*.
4. Sec. 49, *De Motu*.
5. Sec. 51, *De Motu*.
6. Sec. 115, *Principles*.
7. Ibid.
8. Sec. 115, *Principles*, first edition; my emphasis.
9. Sec. 113, *Principles*.
10. Sec. 115, *Principles*.
11. Ibid.
12. Sec. 115, *Principles*, first edition.
13. Ibid.
14. Sec. 10, *De Motu*.
15. Sec. 11, *De Motu*.
16. Ibid.
17. K.P. Winkler, op. cit., 1986, p. 35.
18. Sec. 64, *De Motu*.
19. Ibid.
20. K.P. Winkler, op. cit., 1986, p. 28.
21. Ibid.
22. Ibid.

The Notion of Economy and the Method of its Science According to Lionel Robbins*

RICARDO F. CRESPO
Faculty of Philosophy, Universidad Nacional de Cuyo, Gutiérrez 361
5500 Mendoza, Argentina

The paper at hand argues for the need to revise the epistemological status of Economics. This version should recognize the practical, i.e., moral, dimension of its subject, which is an aspect or a specific type of human action. This act possesses dimensions both technical and ethical which should be integrated into an extensive knowledge, according to the characteristics of the classical practical sciences. This revision would effect many rectifications from which two very clear conclusions would arise. The first one would consist of the recognition of the inaccuracy of Economics' statements. In this regard, important steps have been taken, especially by authors such as George Shackle, Ludwig Lachmann and Jack Wiseman, who are more or less tied to the Austrian School. The second conclusion would be the need to abandon the scientific criterion imposed by Max Weber on the social sciences: the neutrality concerning values—*Wertfreiheit.* Professional economists have almost ignored this latter aspect. However, in spite of the weight of that paradigm, there are some economists who have foreseen the need to include values in Economics.[1] This article centres on the thinking of Lionel Robbins. His conceptual ideas about the economy are centred around the notion of human action. The implicit epistemology of his methodological writings, in turn, evolved during his life time toward a position that tends to give room for the two conclusions of the necessary rectifications mentioned above.

Robbins is a stranger to the readers of philosophical journals. Since Robbins' main contributions belong to the field of the conceptual and methodological definitions, he is not very well-known even by economists, who are often concerned, above all, with more concrete matters. For this reason, I shall first devote a few words to his life.[2]

* I want to thank William J. Baumol, Max Steuer, Peter Koslowski, Heinz Grossekettler and an anonymous referee for useful comments on this paper. I also owe thanks to Ignacio Quiroga, who helped me in the English redaction.

Lionel Charles Robbins was born in Sipson Middlesex, England, near London, on November 22, 1898. From a very young age he was interested in literature; therefore, when his father suggested that he attended the University, he registered in the Faculty of Arts at University College, London. His studies had to be suspended due to his participation in the First World War. Because of his preoccupation with social matters he then entered the London School of Economics and Political Science (LSE) in 1920 and graduated in the autumn of 1923. His first year of work was with Sir William Beveridge, a former Professor. During the following academic year he worked as a Tutor in Oxford (New College) and then he returned to the LSE as a Lecturer. He remained in Cambridge from 1926 to 1927, and in 1929 he was appointed Professor of Economics at the LSE, a position that he held until 1961. His long academic life allowed him to travel and to teach in diverse universities in Europe, America and South Africa. Thus, he met many of the most important economists of this century.

In the 1930s he became a member of the Government Committees of Consulting. During the Second World War he was the Director of the Economic section of the War Cabinet which dealt with economic affairs related to problems occurring during the War and the post-War period. He was a delegate to the Bretton Woods Conferences dealing with the monetary system and international commerce.

Whenever possible he manifested his interest in art, by being involved in the administration of certain institutions, like the National Gallery and the Royal Opera House. During the period from 1954 to 1955 he was President of the Royal Economic Society. In 1960 he was appointed Chairman of the *Financial Times* (until 1970), and he was also named President of the Committee for the Reform of Higher Education in Great Britain. Between 1962 and 1967 he was President of the British Academy, of which he had been a member since 1941. In 1968 he was named *Companion of Honour* and *Life Peer*, and in the same year he also acted as First Chancellor at the University of Stirling. In spite of these commitments, by continuing his educational and managing tasks, he never lost his contact with the LSE, and from 1968 to 1974 he was President of its Directive Council. He died on May 15, 1984, at the age of 85.

Upon shortly reviewing his life one can see that Robbins was a complete man with an extensive academic career. He was able to maintain an equilibrium between his public and private life, all the while cultivating his artistic sensibility, his humanistic interests, as well as devoting himself to his family and friends. In spite of his contributions to economics the totality of the man should not be forgotten. I feel that Robbins' involvement with practical matters has greatly influenced his economic thought, and in this article I shall try to emphasize some aspects of this thought that I consider relevant to the methodological problems of economics.

Robbins' interests were especially theoretical and methodological. Subsequently his interests revolved around political matters and economic history, but he never abandoned his interest in methodology. Robbins' most important book, the *Essay on the Nature and Significance of Economic Science*, deals with the nature and the meaning of economic science.[3] By means of this work, Robbins has become a classic authority concerning the basic statements and definitions of economic science and activity. The Spanish philosopher Antonio Millán Puelles says that, in his opinion, 'the contribution of L. Robbins is exemplary; he is perhaps the most profound economist who has treated the *venatio essentiae* of this knowledge, and undoubtedly the most philosophical one.'[4] The economist William J. Baumol has affirmed that the *Essay* 'can only be regarded as a dangerous revolutionary document. It caused an upheaval in the settled habits of thought of professional economists very soon after its appearance, and it engendered a controversy which has not yet abated-controversy which continues to be creative now as it was then. Even those who disagree most strongly with the positions taken in this book will readily acknowledge that they are substantial, and that they have changed the course of the Economics. . .'[5] Finally, I would like to quote Israel M. Kirzner, who stated: 'Something of a turning point in discussions on the nature of the science and the economic affairs came in 1930 with the appearance of Robbins' *Nature and Significance of Economic Science*. Since the publication of his book, discussions of the problem of definition have invariably tended to resolve around Robbins' definition, or at least to take it as a starting point.'[6]

Robbins' ideas, as they are expressed in the *Essay*, are paradigmatic of modern economic theory. He had the advantage of being at the crossroads of several currents of thought: first, that of classical English Economics, which he understood very well because of his studies of the history of economic thought; second, the deductive methodology of some of his British ancestors; third, the English marginalism of Jevons, Wicksteed and Marshall: Robbins was above all a Neoclassical economist; and finally, the Austrian School, from which he was an important promoter in England.

Robbins always wanted, but was never able, to write a Treatise on Economics. He once wrote three chapters, but then discontinued his writing when he decided that it was more important to his country that he accepted the Chairmanship of the Committee of Reform of the Higher Education. He therefore abandoned this subject, and never resumed the project.[7] When he wrote the *Essay*, he did not intend to do his *opera magna* (he was only 33 years old), but only to put a little order in the methodological questions.[8] During his years as a student at the LSE the fundamental textbook was the *Wealth* by Edwin Cannan which defined the economy in terms of causes of the material welfare.[9] But certain activities such as war and art, observed Robbins, also have economic

aspects and do not correspond to Cannan's categories. The essential relationship with scarcity, recognized Robbins, was not his invention; it was present in the marginal utility theory of the Austrians and in Philip Wicksteed. Robbins only tried to order the ideas using the simplicity offered by the ends-means framework.

The Essay appeared in 1932 and was better received than Robbins had expected, but it also provoked negative comments. The two fundamental criticisms were, on the one hand, the amplitude of the definition, due to its formalism, and on the other hand, its narrowness, due to the fact that it excluded normative elements. Neither of the two worried him because he did not agree with them. He felt that his critics did not thoroughly understand the economic problem. What worried Robbins though was the fact that he found a great deal of 'essentialism' in his conception of economic generalizations. Therefore he stated that in the second edition he put a greater emphasis (but not all that he had wanted), on the need to verify the principles in reality as the next section of this article shall demonstrate.

THE NOTION OF *ECONOMY* ACCORDING TO ROBBINS

Robbins' contributions are various, but I concentrate here on his conceptual and methodological ideas. Robbins' definition of the economy has become famous: 'Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.'[10] The ends are given. The question of their determination is not a matter of economic science since they are an 'irrational' element, external to Economics. This exclusion of the ends from the economic environment consequently implies the exclusion of the values. The means, in turn, are scarce. Without this condition, i.e., scarcity, there is no economic problem. As Kirzner highlights, that is his central concept.[11] From this point of view it is easy to realise why he develops his notion of opportunity cost. 'The conception of costs in modern economic theory, Robbins affirms, is a conception of displaced alternatives: The cost of obtaining anything is what must be surrendered in order to get it. The process of valuation is essentially a process of choice, and costs are the negative aspects of this process. In the theory of exchange, therefore, costs reflect the value of the things surrendered. In the theory of production they reflect also the value of the alternative uses of productive factors—that is, of products which do not come into existence because existing products are preferred.'[12] He simultaneously concludes that the conception of costs as quantities of goods is not acceptable.

In sum, economy is for Robbins a human type of action, which relates those scarce means, of alternative use, to specific ends coming from the outside. Robbins is recognized for his ability to express in just a few words the essence of the economic. In spite of the difficulties with his definition

(these difficulties will be pointed out by various subsequent critics), Robbins was successful in centring the discussion of the economic problem around the notion of human action.

This vision could appear to be an Austrian one. Nevertheless, as Kirzner points out, is different: 'Economizing—Robbins' conception —consists in the allocation of scarce resources among competing ends. Acting, in the praxeological sense, consists in selecting a pattern of conduct designed to further the actor's purposes.'[13]

We must also search Robbins' roots in the classical and neoclassical tradition. Robbins' concern with the classical economists thinking should be remembered.[14]

ECONOMIC METHODOLOGY: FROM AXIOMATIC TO HYPOTHETICAL DEDUCTIVISM

In his methodological and implicit epistemological thought, two problems are distinguishable: first, that of the methodology itself, and second, the epistemological question of the inclusion or exclusion of value judgements in economic science.

Regarding the first problem, we must divide his ideas into two phases. The first one corresponds to the *Essay*, more strictly to its first edition of 1932. In this Robbins followed the deductivist tradition of his ancestors Nassau Senior, John Stuart Mill and John E. Cairnes. Economic laws are necessary deductions of always valid axioms known by external or internal observation. The most important step in the methodology is to clearly state the axioms since all economic science is supported by their validity. When, upon returning to the facts, the axioms do not coincide with the laws, we are faced with a problem of applicability, which is resolved by auxiliar postulates or axioms. Economic science is exact in itself because of the universal validity of its axioms. This is the phase of Robbins' thought about which, as we have already clarified, he had many complaints due to its excessive 'essentialism'. In the second edition of the *Essay* he already tries to give a hypothetical nuance to the axioms. This is more evident in his article 'Live and Dead Issues in the Methodology of Economics' of 1938.[15] In it he states: 'I do not think that there is a single professional economist living who would dispute that the appropriate method of Economics is the construction and development of hypothesis suggested by the study of the reality, and the testing of the applicability of the results by reference back to reality.' There could be differences in the specific kinds of deduction and induction applied by one author or another. 'But on the necessity for the combination of deduction and induction, of theory and realistic studies, there is no dispute among sensible people, adds Robbins. . . We all agree that in general we need both induction and deduction, observation and theoretical system.'[16]

Subsequently, his contact with Karl Popper, a colleague at the LSE, assured him of the correctness of the application of the hypothetical-

deductive method. Various testimonies of Popper's influence still remain. I will begin by quoting the oldest one. In his *Autobiography*, published in 1971, Robbins devotes the chapter seven to his publications and controversies surrounding his works during the years from 1929 to 1939. His most important book during this period as he himself indicates was, of course, the *Essay*. Talking about it, he confirms the majority of the ideas that he already expressed, but he also adds: 'There were, however, sections of the *essay* about which I am far less complacent. In its original form the chapter on the nature of economic generalizations smacked too much of what nowadays is called essentialism. That is to say, it seemed to suggest that the sanction of economic generalizations lay in the conceptions from which they were deduced and these conceptions in some ultimate way were part of the nature of things; and although this puts it crudely, there was indeed something of that in my thinking at that time. In the second edition I tried to put this right and to make it clear that the ultimate assumptions were elementary facts of experience[17] whose appropriateness was always subject to testing by reference back to reality. But I am not at all sure that even then I was successful in achieving a completely convincing formulation. Moreover, while I have very little to retract in what I say in the next chapter about the limited predictive value of time series and suchlike statistical material, I am inclined to think that I did not emphasize sufficiently the indispensability of such operations as means of testing the assumptions of the more complicated theoretical models. This part of my book, more than any other, reflects the circumstances in which it was written. It is a reaction—doubtless overdone—against the ridiculous claims of the institutionalists and of the cruder econometricians and an attempt to persuade Beveridge and his like that theirs simplistic belief in 'letting facts speak for themselves' was all wrong. Moreover it was written before the star of Karl Popper had risen above our horizon. If I had known then of his path-breaking exhibition of scientific method as the attempt to test for falsity to reality the models of the imagination, this part of my book would have been phrased very differently.'[18]

From this point I would like to continue to trace Popper's influence on Robbins' thought. In September of 1974 an extremely exclusive Symposium was held in Naftalion, Greece. Spiro J. Latsis had invited a group of specialists to analyze the adaptability of Imre Lakatos' epistemological proposal, the '*Scientific Research Programs*' (SRP), to both physical sciences and Economics. Among the economists attending were A.W. Coats, Terence Hutchison, Neil de Marchi, Herbert Simon, Axel Leijonhufvud, Mark Blaug, Sir John Hicks, and Latsis himself. The lectures of the economists at the symposium were edited in a volume published by Latsis, *Method and Appraisal in Economics*,[19] extensively reviewed by Robbins in 1979.[20] Robbins said: 'Generalizations about physical or economic phenomena are essentially conjectures capable of

deductive development, but while exposed to refutation by appropriate testing, incapable of final and definitive proof. In this respect I am a convinced Popperean. I do not regard Lakatos' Methodology of Scientific Research Programmes as, in any important respect, contradictory to Popper's main contention. Indeed I would regard his prescriptions, his "hard core" propositions, his heuristics postulates, his "protective belt", and his insistence that broad programmes of research should not be deterred by incidental failure as a distinguished spelling out of implications of the main body of Popper's contribution. I confess that while recognizing, in appropriate places, differences in stylistic emphasis in Popper's exposition, I do not regard these differences as constituting essential divergences of outlook justifying the attribution to the author of inconsistent views within the central contribution. I should certainly regard it as a total misapprehension—which I am sure Lakatos would not have wished to propagate—to regard Popper as a *naive* falsificationist. Therefore personally I regard the broad conceptions of the *Logic of the Scientific Discovery* as generally applicable, whatever particular scientific discipline we practice.'[21] These expressions include, besides the adoption of the Popperean methodology, an interpretation of it and of Lakatos' methodology.

The last quotation is from Robbins' Lecture which was included in the third edition of the *Essay* which was given before 2,000 assistants at the annual Meeting of the American Economic Association (September 5–7, 1980, Denver, Colorado). I would like to point out that Robbins was then almost 82 years old. Therefore, we can consider that these were his last words on this subject. There he affirmed: "Let me say at once that I see no reason for denying to the study of the activities and institutions created by scarcity the title of science. It conforms fundamentally to our conception of science in general: that is to say the formation of hypothesis explaining and (possibly) predicting the outcome of the relationships concerned and the testing of such hypothesis by logic and by observation. This process of testing used to be called verification. But, since this way of putting things may involve an overtone of permanence and non-refutability, it is probably better described, as Karl Popper has taught us, as a search for falsification—those hypotheses which survive the test being regarded as provisionally applicable. I am pretty sure that that all the positive propositions of economics conform to this description.'[22] This sounds undoubtedly highly categorical. With these ideas I shall conclude this first part of the subject dealing with the evolution of Robbins' methodological thinking. Evidently he still isn't the deductivist who thinks that axioms are always valid. Instead he introduces the hypothetical element. However, Robbins does not do this because of the practical character of economics. He is only following Popper. In Popper's theory, the uncertainty is not the contingency of the practical thing but the result of his nominalist gnoseology.

ECONOMIC EPISTEMOLOGY: FROM *ECONOMICS* TO *POLITICAL ECONOMY*

Concerning the second methodological problem, more precisely an epistemological one of economics, i.e., the values inclusive in economic science, a small variation in Robbins' thought can also be observed. Robbins supported the Weberian *Wertfreiheit*, i.e., neutrality in respect to values, as a scientific condition, and because of this, as I have already indicated, he excluded the matter of ends of scientific knowledge. However, Robbins maintained, throughout the years, the need to develop a normative knowledge of economic reality. I shall briefly analyze the evolution of our author's thought concerning this topic.

Robbins first studied economic theory (science), but he quickly affirmed that 'all this was in a very high plan of abstraction.'[23] 'There was another level, however, he adds, on which economic analysis was conjoint with assumptions about the ultimate desirable ends of society which, (. . .) had no less a hold in my attention.'[24] Thus, he began to publish a group of writings about the theory of economic policy. His contact with the classical English economists, which increased throughout the years due to his historic interests, gave him more justification to pursue his new point of view. Precisely what the classics had called political economy was a theoretical study similar to the one proposed by Robbins. As he indicates, 'I adopted the habit of designating appointing such interests by the old-fashioned term political economy to make clear their dependence on judgements of value and distinguish them from pure science. Thus I announced my *Economic Planning and International Order* as 'essentially an essay in what may be called political economy, as distinct from economics in the strict sense of word. It depends upon the technical apparatus of the analytic economics; but it applies this apparatus to the examination of schemes for the realization of aims whose formulation lies outside economics, and it does not abstain from appeal to the probabilities of political practice when such an appeal has seemed relevant.'[25] This idea taken from *Economic Planning* published in 1937 grew stronger as time went by. In 1939 he lectured at the LSE on 'Theory of Economic Politicy.' This constituted the basis of another lecture at Manchester University (1950) and also provided the foundation for his book *The Theory of Economic Policy in English Classical Political Economy* of 1952.[26]

In the preface of his book, *Political Economy: Past and Present*,[27] Robbins refers to the last chapter of *The Theory of Economic Policy*. . ., affirming a continuity between both. In the introductory chapter of this last book we find a synthetic formulation of his concept of Political Economy: 'Political economy in my vocabulary is not scientific economics, a collection of value-free generalizations about the way in which the economic systems functions. It is a discussion of principles of public policy in the economic field: and while it makes appeal to the findings of economic science, it also involves assumptions which, in the nature of things, lie outside

positive science and which are essentially normative in character. It consists of prescription rather than description; although, since it is concerned with practice, its recommendations make use of what aspires to be a scientific examination of the results of action rather than wishful thinking regardless consequences. Political economy thus conceived is essentially a search for solutions to problems of policy.'[28] Robbins continues explaining that the use of the term *Economics* has become the usual name for value-free economic science. In this book he also refers to the first chapter of his *Politics and Economics: Papers in Political Economy* of 1963[29] in which he says that every theory of political economy profoundly depends on judgements and political appraisals. Ignoring the distinction between political economy and value-free economic science, he adds, would be a mistake. But the idea that a system can be built with principles that are consequences of the results of a positive analysis also implies almost the same confusion. Every theory of political economy has to depend, in part, on conventions and appraisals that come from outside the science.[30]

In his already quoted review of Latsis' book (1979), Robbins returned to the theme: political economy should respond to the normative teachings of ethics and social philosophy that involve value judgements.[31] Finally I would also like to mention the article 'Economics and Political Economy' (1981). Since the classical political economy included value judgements, Robbins explained again that this term was left aside and the term *Economics* began to be used. His suggestion here, as in his *Political Economy: Past and Present*, is to revive this term, political economy, in order to emphasize a knowledge that overtly deals with political suppositions and value judgements. Many economic matters correspond to this new and old knowledge. '. . .In the application of Economic Science to problems of policy, Robbins affirms, I urge that we must acknowledge the introduction of assumptions of value essentially incapable of scientific proof.'[32] It is not science, but a 'branch of intellectual activity.'[33] This last phrase proves that Robbins persists in having positivist prejudices. But in spite of these ideas, the recognition of a 'branch of intellectual activity' status for the so-called political economy is a great advancement.

Robbins realizes that a theoretical knowledge concerning economic policy exists; i.e., concerning the application of Economics to concrete matters of social life. The strength of the paradigm that identifies scientificity with value neutrality prevents him from calling it science; but it is at least a type of knowledge. Although opiniative, it is a worthwhile intellectual knowledge. Robbins realizes that Economics cannot be an end in itself. There are only a few economists interested in the strictly pure theoretical knowledge.[34] As Robbins explicitly says, economics is a theoretical study of a practical matter. Economics is directed toward a subsequent application. Economic science (theory, analysis) is not sufficient. However, neither is political science, nor political philosophy

nor ethics. The economic is concerned with specific matters that do not belong to the political field. However, the economist cannot accomplish his practical task only by using his theoretical knowledge. He needs an additional field of knowledge similar to that of the classical English economists, i.e., political economy. Robbins, with his years of experience, sees and indicates the need for this knowledge, and also suggests the studies that should be composed it. Once he said: 'We must be prepared to study not merely economic principles and applied Economics. . . We must study political philosophy. We must study public administration. We must study law. We must study history which, if it gives rules for action, so much enlarges our conception of possibilities. I would say, too, that we must also study the masterpieces of imaginative literature.'[35] In another lecture he insisted on studying political science and economic and general history and he also affirmed: 'I suspect that, in the ideal state, Economics would be taken as a second degree after some short experience of practical life.'[36]

Robbins sets up a bridge between modern and classical Economics. Following Robbins, I suggest a renaissance of the ideas of the classical economists and philosophers in a science that is political economy and that coincides with the characteristics of Aristotle's practical science. In such a practical science economic theory would be a subordinated branch of the latter. The evolution of Robbins' methodological thinking, because of his insistence on hypothetical character of Economics as well as for his theory of an evaluative economic knowledge can be interpreted as aiming at this practical, i.e., moral economic science.

SUMMARY

This paper introduces the life and ideas of the English economist Lionel Robbins. Robbins made a relevant contribution to the definitions of the main concepts of the economy and to the methodology of its science. In the field of theoretical economics he continued not only the line of the English economists, taking into account both marginalist and classical elements, but also that of the Austrian school. His methodological ideas evolved from an axiomatic-deductive position to a hypothetical-deductive one. He also acknowledged the need of a normative knowledge dealing with economic matters, and proposed to give it the old classical name of 'political economy'. Robbins' methodological thought in time is compatible, in part at least, with today's intentions to consider economics as a practical, i.e., moral science.

NOTES AND REFERENCES

1. Cf. my book *La economia como ciencia moral*, Educa, Buenos Aires, 1997, and my papers 'La cuestion metodologica en Keynes padre: un acercamiento a la consideración práctica de la economia', *Philosophica*, 17, Valparaiso, 1995, 'La economia como ciencia práctica', *Sapientia*, Buenos Aires, 1997.
2. The main source of the data that I shall cite has been Robbins', *Autobiography of an Economist*, Macmillan, London, 1971. I have also consulted *The New Encyclopaedia Britannica, Micropaedia*, Vol. 10, fifteenth edition, 1986, p. 102; and Robbins' obituary in *The Economic Journal*, Official Organ of the Royal Economic Society, performed by A.R. Prest, James Meade and Willian J. Baumol, Feb. 1985, pp. 3–7.
3. Macmillan, London, 1932. Since some relevant changes were introduced, the second edition, 1935, is usually used.
4. *Economía y libertad*, Confederación Española de Cajas de Ahorro, Madrid, 1974, p. 126.
5. Foreword to the third edition of the *Essay*, Macmillan, London, 1984, p. vii.
6. *The Economic Point of View*, Sheed and Ward Inc., Kansas City, 1976 (first edition., 1960, pp. 108–9.
7. Cf. ibid., pp. 271 and ss.
8. Cf. ibid., p. 145.
9. Cf. *Wealth: A Brief Explanation of the Causes of Economic Welfare*, 1936.
10. Op. cit., p. 16.
11. Cf. op.cit., p. 110 ff.
12. 'Remarks Upon Certain Aspects of the Theory of Costs', *The Economic Journal*, March 1934, pp. 1–16.
13. Op. cit., p. 161.
14. For example his books *The Theory of Economic Policy in English Classical Economics*, New York: Saint Martin's Press and London: Macmillan, 1952; *Robert Torrens and the Evolution of Classical Economics*, New York–London: St. Martin's Press–Macmillan, 1958. His articles, 'Malhus As an Economist', *The Economic Journal*, 1867, pp. 256–261; 'Pack on Mill', *Economica*, 1957, pp. 250–59. His reviews: to *Jeremy Bentham: An Odyssey of Ideas*, by Mary Mack; to *The Collected Works of John Stuart Mill*; to *The Life of Adam Smith*, by John Rae; to *J.R. McCulloch: A Study in Classic Economics*, by D.P. O'Brien; to *John Rae, Political Economist: An Account of his Life and a Compilation of his Main Writings*, by R. Warren James; to *The Saint Simonians, Mill and Carlyle: A Preface to Modern Thought*, by R. Pankhurst, all in *Economica*.
15. *Economica*, London, August of 1938.
16. 'Live and Dead Issues. . .', op. cit., pp. 346 and 352.
17. Here he uses this expression, 'elementary facts of experience'; in the next two quotations he will use the terms 'hypothesis' and 'conjectures' that appear to be more compatible with his new position.
18. *Autobiography of an Economist*, op. cit., pp. 149–150.
19. Cambridge University Press, Cambridge, 1976.
20. 'On Latsis' Method and Appraisal in Economics: A Review Essay', *Journal of Economic Literature*, Vol. XVII (Sept./79), pp. 996–1004.
21. Op. cit., p. 999.
22. Op. cit., p. 2.
23. *Autobiography. . .*, op. cit., p. 150.
24. Ibid.
25. Ibid., and *The Economic Planning and the International Order*, Preface.
26. Macmillan, London, 1952. Translated into Spanish as *Teoría de la política* económica, Rialp, Madrid, 1966. Cf. Preface, pp. 11–13.
27. Macmillan, London, 1976. Based on some lectures given at Cape Town University South Africa, 1974.

28. Ibid., p. 3.
29. Macmillan, London and St. Martin's Press, New York, 1961, p. 3 (nt.)
30. Cf. Ibid., 1965, p. 20.
31. Cf. 'On Latsis'. . .', op. cit., pp. 998–9.
32. Op. cit., p. 9.
33. Ibid.
34. *Political and economy. . .*, op. cit., p. 7.
35. 'The Economist in the Twentieth Century', *Economica*, May, 1949. Also published in *The Economist in the Twentieth Century and Other Lectures in Political Economy*, Macmillan, London, 1956, p. 17.
36. 'The Teaching of Economics in Schools and Universities', *The Economic Journal*, December 1955, cf. pp. 582 and 587.

The Marxian Critique of Justice and Rights: Some Reflections

ALOK TANDON
Tandon Nivas, 53 Ashraf Tola, Hardoi 241 001, Uttar Pradesh

In this paper I am concerned with assessing Marx's critique of justice and rights. But before it is done, his views must be articulated in all its complexity, more so because there is a controversy among scholars regarding Marx's position on these issues since he never specifically spelled it. The least controversial about Marx's views on justice and rights is that judicial conceptions cannot serve as the most explanatory concepts. It is the analysis of the mode of production that provides the key to understanding the society as a whole because he believes that these basic productive forces and processes are the real foundations on which rises a legal and political superstructure.

I

A lively debate has recently flourished on this question—is the relation between the capitalist and the worker in capitalism unjust in the eyes of Marx. It is primarily focused on Marx's critique of the critical role of judicial conceptions used to make fundamental assessments of a given society. All the four logically possible positions on the issue have been fairly well defended by the scholars, namely:

(1) The relation between capitalist and worker was just (Robert Tucker, 1969, and Allen Wood, 1972, pp. 79, 91).
(2) It was unjust (Husami, 1979; Cohen, 1983, and others; Elster, 1985).
(3) It was just in one respect and unjust in another (Gary Young, 1975–76, 1981).
(4) It was neither just nor unjust (Richard Miller, 1984).

Since all the four positions of Marx considered above are supported by the textual evidence, it may be argued (Cohen and Elster) that Marx just failed to understand his own view of justice as he did not adequately reflect upon it. But this is no reason to feel shy of further explicating the problem. In my opinion, failure to make distinction between Marx's

internal and external criticism of capitalism and inability to distinguish Marx's critique of distributive justice from his critique of civil and political justice are behind much controversy.

With the help of a number of very telling passages (Marx, 1845, pp. 46, 65, 75, 79–80) Tucker-Wood thesis claims that Marx does not condemn capitalism as a system which violates some standard of distributive justice. His argument is as follows: since a standard of justice, according to Marx, can only be meaningfully applied to that mode of production from which it arises and to which it corresponds, the wage relation between workers and capitalist is just according to the only standard of justice which applies to it, namely the standard which requires that equivalents be exchanged for equivalents. Hence Wood concludes that, for Marx, though capitalist exploitation 'alienates, dehumanizes and degrades wage labourers', it is not unjust (Wood, 1979, p. 108).

Several important objections to Wood's interpretation can be raised. As Nancy Holmstrom has pointed out that Wood views the exchange between the worker and capitalist too narrowly, abstracted from its background. Marx's point is that even if the transaction is an exchange of equivalents, it is not a free exchange; the worker is forced to sell his labour power because capitalist controls the means of production. Once we understand the coercive background of the transaction, this cannot be taken to be just (Holmstrom, 1977, pp. 366–67).

Holmstrom's criticism of Wood is plausible one and points to what can be called Marx's *Internal* criticism of certain notions of justice endemic to capitalism. So long as the principle that exchanges are just if and only if they are free agreements among equals is applied myopically, the wage relation appears to be just. But once we take seriously and apply consistently the ideals of freedom and equality and refuse to narrow their application arbitrarily, capitalism's own standard of justice provides material for a critique of capitalism itself (Marx, *Das Kapital*, Vol. 1, p. 271) But, we must explain why, in spite of his frequent use of internal criticism, Marx did not primarily criticize capitalism for its distributive injustices. This leads to *external* criticism of distributive justice of capitalism by Marx.

External criticism is external to the conception of justice under criticism. However, the external perspective, may be either an alternative conception of justice or some non-judicial evaluative conception. Holmstrom is of the opinion that some of the external criticism of capitalism launched by Marx is from the perspective of a communist conception of justice based on communist mode of production. But, as Buchanan has successfully argued, at least in the case of distributive justice, this is false. Marx not only ridicules those critics of capitalism who speak of socialist justice, he seems deliberately to avoid the use of the language of justice and rights himself and advises socialists to cease their preoccupation with such 'ideological nonsense'. Since, for Marx,

distribution in any society is a derivative phenomenon determined by that society's basic productive processes, the root defects of capitalism lie in its basic productive processes and the fundamental superiority of communism will be found in its distinctive productive processes than in its distributive arrangements. The communist mode of production will so reduce the problems of scarcity and conflict that principles of distributive justice will no longer be needed. What Rawls calls the circumstances of distributive justice will either no longer exist or so deminished that they no longer play a significant role in social life. Communist mode of production, according to Marx, not only unfetter the tremendous productivity latent in capitalism, it will also transform the individual into a cooperative, communal being. Man in communist society will find work intrinsically satisfying and his need for human association will be greater than his need for things. Thus, we can say that the basis of Marx's most radical criticism of capitalism and its juridical concepts is his evaluative conception of fully developed communist man. For Marx, the very need for principles of distributive justice is conclusive evidence of defects of productive process that form a society's core. His claim is not that a broader notion of communist productive-distributive justice is needed to solve the problems of distribution, but he points out that once new communist productive arrangements appear, there will be no need for principles of justice for production or distribution.

We can conclude our discussion about Marx's views on justice in the words of Steven Lukes (1985, p. 59) 'What Marx offers is a multi-perspectival analysis in which capitalism's self-justifications are portrayed, undermined from within, and criticized from without, and then both justification and criticism are in turn criticized from a standpoint that is held to be beyond justice.'

II

So far our analysis focused on Marx's critique of distributive justice. Now we shall examine Marx's criticism of the justice of civil and political rights. The most important source of these criticisms is Marx's essay, *on the Jewish Question.* In it, Marx distinguishes between (1) human rights, (2) the rights of man, and (3) the rights of the citizen. Human rights is the most general category which includes the other two as sub-categories. The rights of the citizen are rights of political participation, especially the right to vote. The rights of man include:

(a) The right to freedom of expression, thought and belief (especially religious belief);
(b) The right to equality before the law (the rights of due process);
(c) The right to private property;

(d) The right to security (the right to freedom of the person–protection of life and limb); and
(e) The right to liberty.

The rights of man and of citizen are those rights by which individuals becomes politically emancipated which is granted by the modern liberal state. The main thrust of Marx's argument is that political emancipation falls short of genuine human emancipation which Marx believes can only be achieved in communism. He saw rights: 'nothing but right of a member of civil society, i.e., the rights of egoistic man, of man separated from other man and from the community (Marx, 1943, p. 162). For him securing these rights of 'self interest' of a monadic and 'restricted individual' is basically reducing it to the protection of private property and having an illusion that there is a separate sphere of political emancipation as a surrogate for and precursor of general human emancipation. His point seems to be that existence of and the very need for the right to property—both marks and perpetuates a situation in which there are basic conflicts of interests between individuals. In such circumstances, other individuals are perceived at best as mere means and at worst as lethal threats. And if individuals would not so view each other but for the existence of private property, then it is plausible to say that the right to private property is valuable only for egoistic, monadic man. Further, it can be said that right to property separates a person from the community in the sense that it protects him against intrusions by others and absolves him of any responsibility for the welfare of others. But, this does not mean that Marx advocates the replacement of a defective capitalist right to property with a superior communist right to property. His contention is that in communism the sources of conflict will be so diminished that there will be no need for a system of rights to guarantee the individual's freedom to enjoy his share of the social product or to guarantee him a share of control over the means of production. Marx seems to believe that any society in which the potential for interpersonal conflict is serious enough to warrant the establishment of rights to serve as limits on conflict is a deeply defective society. Only on this basis, Marx's scorn for rights in general can be understood.

Marx's criticism of the rights of man, to some extent, carries over to the rights of citizen. There are perfectly legal ways in which differences in wealth and status produce inequalities in the effectiveness with which different individuals can exercise their equal citizenship rights. For Marx, the rights of man and the rights of the citizen are correlatives which mark a division between man's existence as an independent egoist in civil society and his realized life as a citizen, a moral agent concerned with the common good rather than his own narrow self-interest (Marx, 1943). The implication is that in communism, where the concept of the egoistic, isolated individual is no longer applicable, the correlative concept of

man as citizen, along with the notion of rights of the citizen, will also no longer apply. This does not mean that persons in a communist society will not enjoy free speech or freedom from assault on their lives, it only implies that in communism, there will be no need to guarantee these freedoms.

As in case of Marx's criticism of distributive justice, a distinction can be drawn between internal and external criticism of civil and political rights. His main internal criticism is that so long as wealth can corrupt legal and political processes and so long as differences in social position create inequalities in effectiveness with which different individuals can exercise their equal rights, political emancipation fails to live upto its own ideal of equal citizenship. Marx's external critique is more radical than his internal critique and is launched from a perspective of a society that is beyond the circumstances of rights. He seems to attack the very notion of (legal) right as being an artifact of defective mode of production and believes that such a notion will become obsolete when a superior mode of production comes into being.

III

On the basis of Marx's internal critique of justice, a question can be asked, is not the quest for justice futile because the circumstances of justice are just those conditions in which demands of justice can never be met. The answer is that Marx's theory of class conflict when taken together with his theory of the state, implies that the generic concept of justice makes demands that cannot be met in those circumstances which give rise to that concept and to which it applies. So long as capitalist remains a capitalist and worker remains a worker, the principle of free and equal exchange between the buyer and seller of labour power is not only unsatisfied but unsatisfiable. Moreover Marx's conception of proletarian revolutionary action does not allow any significant motivational role for principles of justice. All that is required, according to him, is the motivation of self or class interest. Marx's scathing criticism of moralizing socialists, like Proudhon, is a case in point.

Now we can understand in what sense Marx thought talk about justice and rights was obsolete verbal rubbish. First, his understanding of social change led him to deny any fundamental explanatory role to juridical notions in general. Second, though in his internal criticism of capitalism he uses bourgeois notions of justice and rights, yet they play no major role in his most radical attack on capitalism from an external perspective. Third, for him, the normative conceptions of communist justice and rights are not needed either to motivate revolutionary action or as basic principles of a new society to which struggle is directed.

Thus we can conclude that Marx's internal and external criticisms are related. His internal criticism plays a destructive role to demolish

arguments of bourgeois ideologist's attempt to show that capitalism is just while his external criticism leads to main thesis that rights will not be needed in communism. This does not preclude the possibility that they will be useful in some stages of transitional period prior to communism.

IV

When we critically examine the views of Marx, it can be said that his was a narrow and impoverished view of the meaning of rights of man. He treated them only as symptomatic of the individualism and contradictions of bourgeois life. He failed to consider their positive, world—historical significance and their applicability to non-egoistic, non-bourgeois forms of social life. Thus, he missed their consequent relevance to the struggle for socialism. Moreover, his thesis that it is only the conflicting class interests that render rights necessary is hard to believe in the light of historical facts and analysis. As Buchanan has convincingly argued (Buchanan, 1982), the protections and guarantees may be required in, at least, the following circumstances: where minorities are disfavoured by democratic procedures, where paternalist policies interfere with individual liberties, where disagreements exist about what constitutes welfare or the common good, where coercion is required for the provision of public goods, and where guidelines and limits must be set to the provision for future generations. We cannot assume that under communism, free from class conflicts, such circumstances as above, would not prevail. Even in an altruistic society, protections from spurious claims will be required to test such claims. Thus, we can sympathize with Steven Lukes (1985, p. 65) who has criticized Marx's views for double narrowness: 'too narrow on account of the significance of rights and too narrow a view of circumstances that render them necessary. The former narrowness has made them seem unimportant, the latter potentially dispensable.'

Buchanan (1982, p. 177) has ably pointed out that 'key elements of Marx's critique of judicial concepts are inadequately supported due to deficiencies in his theory of revolutionary motivation, on the one hand, and his lack of a theory of non-juridical, non-coercive, but highly productive and harmonious social coordination on the other.' Marx's foundational criticism presupposes such a theory but his writings do not supply it. By Marx's interest theory of motivation it is difficult to explain, but equally difficult to ignore the revolutionary's use of violence against members of the proletariat and his reliance upon what Marx called 'obsolete verbal rubbish' about justice and rights.

In spite of the deficiencies discussed above, Marx's critique of juridical concepts is of great theoretical and practical importance. It offers nothing less than a systematic attack on two doctrines which we may be tempted to view as self-evident the principle that the justice is the first virtue of the social institutions and the principle that respect for persons

as right bearers is the first virtue of the individuals. Even if we ultimately reject much of what Marx has to say against these theories, his critique leads us to better understanding of the nature and value of juridical concepts. By invoking rights we acknowledge the inevitability of conflicts without examining their sources. Even if an extensive juridical framework is necessary, we should not assume that it will be adequate for all human relations nor that it can stand alone without the support of deeper effective structures which are not themselves informed by conceptions of rights or justice.

After Marx, no theory of distributive justice can ignore the interdependence of distribution and production. No reform can focus exclusively on distribution to be taken seriously. Similarly, Marx's historicist criticisms make it impossible for the political theorist to make claims about 'universal rights' or 'eternal' principles of justice.

Marx's brilliant criticism of political emancipation is still very important. Any serious normative political theory must face his challenge to basic liberal assumptions that political equality can coexist with socio-economic inequality. We must go beyond equal rights towards effective equality.

And finally, by executing a systematic critique of capitalism without relying primarily on conceptions of rights and justice and by articulating the ideal of a free and humane society that is essentially non-juridical, Marx presents the most radical challenge to the conceptual framework of traditional moral and political theory. Appeals to the 'sense of justice' will not suffice unless they are accompanied by an empirically supported theory of social change.

References

Buchanan, A.E., *Marx and Justice: The Radical Critique of Liberalism*, Methuen, London, 1982.

Cohen, M., Nagel, T., and Scanlon, T., *Marx, Justice and History*, Princeton University Press. Princeton, New Jersey, 1980.

Elster, J., *Making Sense of Marx*, Cambridge University Press, 1985.

Holmstrom, N., 'Exploitation', *Canadian Journal of Philosophy*, Vol. 2, 1977.

Husami, Z.I., 'Marx on Distributive Justice', *Philosophy and Public Affairs*, 8.1 (Fall), 1978, reprinted in Cohen, Nagel and Scanlon, 1980.

Lukes, Steven, *Marxism and Morality*, Oxford University Press, New York, 1987.

Marx, K., *Capital*, Vol. II, 1861–79, Foreign Language Publishing House, Moscow, 1957.

Marx, K., *Capital*, Vol. III, Foreign Language Publishing House, Moscow, 1962.

Marx, K., *Wages, Price and Profit*, S.W.1., 1865.

Marx, K., *Critique of Grotha Programme*, S.W.2, 1875.

Marx, K., *Notes on Adolph Wagner*, 1879–80. Translation in Terrell Carier (ed.) *Karl Marx: Texts on Method*, Blackwell, Oxford, 1975.

Marx, K., *On the Jewish Question*, MECW 3, 1843.

Marx, K., *The Holy Family* MECW 4, 1845.

Marx, K., *The German Ideology*, MECW 5,1845–46.

Miller, R., *Analyzing Marx*, Princeton University Press, Princeton, 1984.

Tucker, R.C. (1969), *The Marxian Revolutionary Idea*, W.W. Norton, New York,1969.

Wood, A., 'The Marxian Critique of Justice', *Philosophy and Public Affairs*, 1, 3, (Spring), 1972, reprinted in Cohen, Nagel and Scanlon, 1980.
Wood, A., 'Marx on Right and Justice: A Reply to Husami, *Philosophy and Public Affairs*, 8, 3, (Spring), 1979, reprinted in Ibid.
Wood, A., *Karl Marx*, Routledge and Kegan Paul, London, 1981.
Young, G., 'The Fundamental Contradiction of Capitalist Production', *Philosophy and Public Affairs*, 5, 1975–76.
Young, G., 'Justice and Capitalist Production: Marx and Bourgeois Ideology', *Canadian Journal of Philosophy*, 8, 1978.
Young, G., *Doing Marx Justice, Canadian Journal of Philosophy*, Supplementary Vol. 7.

Wonhyo as a Creative Philosopher in Korean Buddhism

SUN KEUN KIM
Dept. of Philosophy, Arsandey, P.O. Boreya 834 007, Ranchi, Bihar

I. THE SIGNIFICANCE OF DHARMA IN BUDDHISM

Although born on Indian soil, Buddhism is one of the three great world religions, with a following which, at the height of its influence, included between a third and a fourth part of the human race, and which even now is not negligible. Its founder, Gautama the Buddha, is universally recognized as the perfect embodiment of the ethical and spiritual ideals which he spent his life proclaiming. Indeed, it is a majestic figure of the Enlightened One, seated cross-legged in meditation. Its influence upon other systems of belief has been profound. Besides being a carrier of culture and civilization for the whole of Asia Buddhism affords the unique spectacle of a doctrine of salvation. In the course of its long history it has given the world, and continues to give, an ethics based on the ideal of absolute altruism. In fact Buddhism is no longer regarded as something to be studied by the professional orientalist but as a way of life for the spiritually committed individual. So I would like to examine the significance of Dharma and briefly survey the evolution of the main schools of Buddhism in outlook of Dharma.

As Mrs Rhys Davids writes, 'the Dhamma has in the history of Buddhism its history'.[1] Stcherbatsky also observes: 'the conception of a *dharma* is the central point of the Buddhist doctrine. In the light of this conception Buddhism discloses itself as a metaphysical theory developed out of one fundamental principle.[2] '*Dharma* is derived from the Sanskrit root *dhṛ* which means to hold together'. As P.T. Raju said, 'Of the Sanskrit word *Dharma* as used in Buddhistic philosophy, we might say the same thing which has been said of its Latin equivalent 'res', namely that is a blank cheque which has to be filled in accordance with the exigencies of the context. 'Dharma' means, in Buddhistic Sanskrit, law, rule, faith, religion, world, phenomena, thing, state, etc.[3]. Thus the term has a variety of meanings. Most narrowly interpreted it means simply the teaching of the Buddha. But in its widest sense it also reaches out to include cosmic order. While expounding this metaphysical significance of *Dharma* in Buddhism, P.T. Raju contends that *Dharma* is to be understood as the

ideal reality towards which the whole universe should move.[4] The ideal reality is the ideal not only of man's life but of the whole nature. It is the principle of righteousness which exists not only in a man's heart and mind, but in the universe also.

II. THE EVOLUTION OF THE MAIN SCHOOLS OF BUDDHISM IN THE OUTLOOK OF DHARMA

The teaching of Buddha were oral and were recorded much later by his disciples.Buddha was primarily an ethical teacher and a social reformer than a theoretical philosopher. He referred to a number of metaphysical views prevalent in his times and condemned them as futile. Whenever metaphysical questions were put to him, he avoided them saying that they were neither profitable nor conducive to the highest good.[5] He repeatedly told his disciples:

'Two things only, my disciples, do I teach-misery and the cessation of misery. Human existence is full of misery and pain.Our immediate duty, therefore, is to get rid of this misery and pain. If instead we bother about barren metaphysical speculations, we behave like that foolish man whose heart is pierced by a poisonous arrow and who, instead of taking it out whiles away his time on idle speculation about the origin, the size, the metal, the maker and the shooter of the arrow.'[6]

Troubled by the sight of sickness, decay and death, Buddha left his home to find a solution of the misery of earthly life. *Pratītyasamutpāda* is the solution he found. Buddha identifies it with the Bodhi, the enlightenment which dawned upon him under the shade of the Bodhi-tree at Bodhgaya and which transformed the mortal Siddhartha into the immortal Buddha. He also identifies it with the *Dharma*, the Law: 'He who sees the *Pratītyasamutpāda* sees the *Dharma*, and he who sees the *Dharma* sees the *Pratītyasamutpāda*.[7] Failure to grasp it is the cause of misery. Its knowledge leads to the cessation of misery. Nāgārjuna salutes Buddha as the best among the teachers, who taught the blessed doctrine of *Pratītyasamutpāda* which leads to the cessation of plurality and to bliss.[8] The doctrine of *Pratītyasamutpāda* or Dependent Origination is the foundation of all the teachings of the Buddha and other teachings can be easily deduced from it as corollaries. Buddha says the definition of *Pratītyasamutpāda* as follows: 'that being present, this becomes: from the arising of that, this arises. That being absent, this does not become; from the cessation of that this ceases[9]. The doctrine of *Pratītyasamutpāda* is the chief contribution of Buddhism to Indian thought. It is contained in the second Noble Truth which gives us the cause of suffering, and the Third Noble Truth which shows the cessation of suffering. Thus Buddha preached the truth he had discovered, without distinction of caste, creed or colour. So Buddhism was embraced by the poor and the rich, the low and the high, the dull and the intellectual alike.

The Dhammapada of Buddhism gives a rational interpretation of the word *Dharma* (Pali, *dhamma*) in ethical and spiritual terms. Buddha says, 'Not by matted hair, not be lineage, not by caste does one become a Brahmin. He is a Brahmin in whom there are truth and dhamma. He is blessed.'[10] Underneath the Sāla trees at Kuśinagara, in his last words to his disciples, the Buddha said: 'Make of yourself a light. Rely upon yourself; do not depend upon any one else. Make my *Dharma* your light Rely upon my teaching.'[11] And he further says, 'The teachings which I have given you, I gained by following the path myself. You should follow these teachings and conform to their spirit on every occasion. If you neglect them it means that you have never really met me.It means that you are far from me, even though you are actually with me, but if you accept and practise my teachings, then you are very near to me, even though you are far away. . . . But the true Buddha is not a human body: it is Enlightenment. A human body must vanish, but the wisdom of Enlightenment will exist for ever in the truth of the *Dharma* and in the practice of the *Dharma*. He who sees merely my body does not truly see me. Only he who accepts my teaching truly sees me. After my death, the *Dharma* shall be your teacher. Follow the *Dharma* and you will be true to me.'[12] These teachings are the central contents of *Dharma* of early Buddhism.

The most important point for us to note in understanding the Buddhistic view of reality is the *Dharmakāya* of Buddha. The Hīnayānists thought of Buddha as only a historical person, who taught a noble doctrine. A school called the Lokottaravādins maintained that the Body of Buddha was supramundane, and that made of flesh and blood was not Buddha's real body. As P.T. Raju observes, 'the idea that Buddha's body could not be human led to the speculation about its actual nature. If it is divine, it must be the essential nature of the world, its *Dharma*. Then was formed the concept of the *Dharmakāya*.'[13] The Mahāsanghikas held that Buddha's body pervaded all directions, and the apparent body was not the real body. The Mahāyāna took over these ideas and developed them further, Buddha had several bodies. There is the physical body born to his parents. Then there is the body of doctrine (*Dharmakāya*) which he taught to his followers.[14]

Buddha's *Dharmakāya* is treated by Aśvaghośa as a metaphysical entity. Every man, it was declared in the Mahāyāna, could become a Buddha through discipline and knowledge. If *nirvāṇa* is the ideal of every life, then to enter *nirvāṇa* means to become one with the *Dharmakāya* of Buddha. The real nature of Buddha's body is *Dharma* or Law, to become Buddha is to become his *Dharmakāya*. To become a Buddha means to know the *Dharma* or Law. So the *Dharmakāya* is the knowledge of the oneness as realized at the Bodhi.

For *Āśvaghośa, nirvāṇa* and the *Dharmakāya* are the *Tathatā* (Suchness). The account of the Mahāyāna schools may be begun with that of the

Bhūtatatathatā school of Aśvaghoṣa. The doctrine of Bhūtatatathatā, Sogen tells us, is explicitly explained in Aśvaghosa's *Mahāyāna-śraddhotpādaśāstra*, which is known in translation as *The Awakening of Faith*. Suzuki says that the Sanskrit original of this work seems to be lost, and we get it only in translation.[15]

Suzuki wrote: 'the three points constituting the gist of the Śāstra are: (1) the conception of Suchness (*Bhūtatatathatā*); (2) the theory of the triple personality; (3) the salvation by faith or the *Sukhavati* doctrine.[16]

The Suchness (*Bhūtatatathatā*) is not only the essence but also the source of the world. It exists 'in all things, remains unchanged in the pure as well as in the defiled, is always one and the same (*samatā*), neither increases nor decreases, and is void of distinction'.[17]

Yamakami Sogen wrote: 'The Suchness (existence as such), *Bhūtatatathatā* is called by as many different names as there are phases of its manifestation. It is *nirvāṇa* when it brings aboslute peace to a heart egoistic and afflicted with conflicting passions; it is *Bodhi* or perfect wisdom, when we regard it as the source of intelligence; it is *Dharmakāya*, when it is called fountainhead of love and wisdom; it is *Kuśalamūlam*, or the *summum bonum* when its ethical phase is emphasized. . . . It is the *Tathāgatagarbha* or the Womb of *Tathāgata* when the analogy from Mother Earth (where all the germs of life are stored, and all precious stones and metal are concealed under the cover of filth) is drawn.'[18] *Tathatā* (Suchness), which is the highest *Sāmanya* or universal, is the only reality, and all particulars, which are forms of individuation, are only appearance.[19] Truth is beyond all distinctions; even the distinction between the knower and the known, the speaker and the spoken, does not obtain in it.[20] Again, there is a two-fold aspect in suchness if viewed from the point of its explicability. The first is trueness as negation (*śūnyata*), in the sense that it is completely set apart from the attributes of all things unreal, that is the real reality. The second is trueness as affirmation (*aśūnyata*), in the sense that it contains infinite merits, that it is self-existent.[21] It is neither unity nor difference, nor both nor neither.[22] To sum up, the concept of *Tathatā* (Suchness) is regarded as containing all perfect attributes.

The *Prajñāpāramitās* constitute the first Mahāyāna literature that is known. If we compare Buddha's original sayings to the Upaniṣads and the Mahāyāna systems to the *Vedāntic* systems, then the *Prajñāpāramitās* may be reasonably compared to the *Brahmasūtras*. Nāgārjuna wrote a work called *Mahāprajñāpāramitāśāstra*, and Maitreyanātha composed a number of *Kārikās*. So the *Prajñāpāramitā* literature seems to be commented upon both by the Madhyamikas and the Yogacārins; just as the Brahmasūtras are commented upon by the different Vedāntic schools.[23]

The major theme of the *Prajñāpāramitas is śūnyata* (emptiness) and *niḥsvabhāvatā* (naturelessness). Haribhadra repeatedly quotes that what is called *śunyata* is the same as *Pratītyasamutpāda* (dependent emergence),

an idea which has been clearly demonstrated by Nāgārjuna.[24] The reason given here is that there is no *Dharma* (entity), which is devoid of the *Dharmadhātu*.[25] This is of course the final reason, because the *Dharmadhātu* is the same as the *śūnya*. As determinations like one and many are not real, things have no nature (*svabhāva*).[26] To have a nature means to be characterized, and every character is a determination. But determinations are not real, and so things have no *svabhāva* (self nature). Reality is beyond thought.[27] True existence is inexpressible; it is beyond all determinations. As it is the *svabhāva* or own nature of everything, it is the law of all. In the *Mahaprajñāpāramitā Hṛdaya Sūtra*, we read: 'Emptiness is not different from form, form is not different from emptiness. What is form that is emptiness, what is emptiness that is form. Thus, perception, name, conception, and knowledge also are emptiness.[28] To sum up, the final nature of everything is the *śūnya*.

The Mādhyamika or the Śūnyavāda of Nāgārjuna is a direct result of *Prajñāpāramitās*. For Nāgārjuna, everything is pure *Śūnya*, which is the same as the *Tathatā* or *Dharmakāya*. For the Mādhyamika, for whom *nirvāṇa* is the *śūnya*, the *Dharmakāya* also is the *Śūnya*. For the Vijñānavāda, everything is pure *Vijñāna*, which the same as the *Tathatā* or *Dharmakāya*. It is generally said that the Mādhyamika's theory is logical or epistemological while that of the Vijñānavāda is psychological.

As mentioned above, we have briefly examined the evolution of the main schools of Buddhism in outlook of *Dharma*. To sum up, the *Dharma* of the Buddha means truth, that which really is. It also means law, the law which exists in a man's heart.

III. THE EVOLUTION OF BUDDHISM IN KOREA

In the course of its long history it spread like wild fire far and wide from the lofty Himalayas to Cape Comorin and ranged beyond the frontiers of its homeland to Srilanka, Myanmar, Thailand, Cambodia, Laos and then again to Nepal, Tibet, Mongolia, China, Korea and Japan. In fact, it became a world-religion and a great cultural force, at least in Asia.

By the Hīnayāna is meant that form of Buddhism which prevails in Srilanka, Myanmar, Thailand, Cambodia and Laos, the literary basis of which is the texts of the Pāli Canon.

By the Mahāyāna is meant the form of Buddhism flourishing in Nepal, Tibet, Mongolia, China, Korea, Japan and parts of Central Asia, its literary basis being the voluminous texts of the Chinese and Tibet Sacred Canons, the bulk of which are translations from Sanskrit Originals.

Buddhism was introduced into Korea early in the era of the Three Kingdoms. According to Korean history, there were originally three kingdoms in ancient Korea. Shilla was established by King Hyock-Keo-Se Park in 57 BC. in the southern part of the Korean peninsula. The capital city of Shilla was Kyong-Ju. This city still exists today in modern Korea.

The Paeckche dynasty was established by King Onjo in 18 BC. Its territory was located near modern Seoul and Chung-Chong Do province in the central part of the Korean peninsula. Its capitals were Booyo and Kong-Ju. These cities still exist today in modern Korea. The Koguryo Dynasty was established by King Dong-Myong in 37 BC. Its capital city was Pyong-Yang. Its territory included the northern part of the Korean peninsula and part of what is now Manchuria. Pyong-Yang is the capital of North Korea. These three kingdoms of Shilla, Paeckche, and Koguryo were eventually united into one kingdom by the Shilla King, Moo-Yul in AD 668. This Shilla Dynasty ended in AD 935 and was followed by the Koryo Dynasty (AD 936–1392). Among the three kingdoms, Koguryo was the first to accept Buddhism. In AD 372, the second year of the reign of King Sosurim, a Buddhist monk named Sundo, who had been sent by the Chinese Former Chin.

King Fu Chien arrived in Korea bringing Buddhist scriptures and religious images. King Sosurim welcomed him and later erected temple for the monk. This is the origin of Korean Buddhism.[29]

Thirteen years later, in AD 384, the first year of the reign of King Chimryu of the kingdom of Paeckche, Buddhism was introduced to Paeckche by Malananta, an Indian monk who came to Paeckche through the Eastern Chin kingdom in China. Buddhism was officially recognized in Shilla, the kingdom located on the southern part of the Korean Peninsula, in AD 527 (the fourteenth year of the reign of King Bubhung. That was the year that Lee Cha Don was martyred for the Buddhist faith in Shilla.

Buddhism became very popular in all of the three kingdoms. However, the golden age of the faith began after unification of the kingdoms by Shilla in the middle of the seventh century (AD 668), even though the Buddhist culture of Koguryo and Paeckche faded away after this unification.

Buddhism has not only influenced the spiritual history of the nation but has also exercised great influence in all branches of Korean politics, economics, society, and culture. This is evidenced by the fact that Buddhism has had the greatest influence on the Korean cultural heritage.

As Japanese Buddhism is a cultural product of Japan, and therefore different from Chinese Buddhism, so is Korean Buddhism a unique expression of Korean culture. When Indian Buddhism entered Tibet and encountered the native Bon religion, it was transformed into Lamaism, the Buddhism of Tibet. Similarly, in China, when the Indian Buddhist philosophy of Prajña-Śūnyata met Taoism, the result was Ch'an Buddhism, a cultural product of China.

Buddhist doctrines, both Therāvāda and Mahāyāna, originated in India. Those doctrines were expounded in detail by the Chinese and formed the basis for the development of various denominations in China. However, it was in Korea that these denominations were unified

into one Chogye Order of Korean Buddhism. The history of Korean Buddhism, accordingly, is the history of a continuing effort toward this synthesis.

IV. WONHYO AS A CREATIVE PHILOSOPHER IN KOREAN BUDDHISM

Wonhyo (AD 617–686) of Shilla mastered the tenets of various denominations and scriptures and established the foundation for a unifying theory of one Buddhism. Later there was Dae Gak Guksa of Koryo whose ideal of Kyo Kwan Kyom Su, parallel practice of doctrine and seon(Zen), which was based on Tien Tai philosophy, was the unifying force between the doctrine and seon sects. However, it was National Teacher Bojo (AD 1158–1210) of Koryo who finally achieved harmony between the doctrinal and seon sects. This spirit of harmony has become the great tenet of Chogyechong, the unified monastic order of Korean Buddhism. No other Buddhism in the world—India, China, Japan, etc.—enjoys the unity of Philosophy found in Korea. It is important to be familiar with the development of this unique reconcilable character to understand Korean Buddhism.

Since Wonhyo (AD 617–686) played such an important role in establishing the philosophical foundation for a unifying theory of one Buddhism, a study of his philosophy and its influence in achieving a unified, non-denominational Korean Buddhism is essential.

The purpose of this researcher is to examine Wonhyo's (AD 617–686) philosophy on the theory of reconciliation of Buddhist doctrines and its influence on selected subsequent Asian thinkers.

The earliest extant record of Wonhyo's central philosophy is found in the Koseon Sa Temple inscription. The inscription reads in part:

> When the Tathgatha was in the world, everybody relied on his perfect teaching. After Buddha's death, however, people's opinions were like showers and pointless theories like rising clouds. Some said, 'I am right: other are wrong.' Others argued, 'mine is like this (but) others' are not like this.' Finally, (theories and opinions) became a flood, . . . the attitude of staying in a deep valley while (avoiding) great mountains or loving emptiness while hating existence is just like the attitude of going into a forest w hile avoiding trees. But one should be aware of the fact that green and blue are identical in essence, and ice and water are identical in origin: a mirror reflects myriad forms, and parted waters will perfectly comingle (once they are brought back together). Thus, he wrote for the book described as a *Treatise on the Reconciliation of All Disputes in Ten Aspects* which everyone accepts. Every body says it is wonderful.[30]

He really had such a far-sighted and macroscopic point of view. He was called the greatest patriarch who was the founder of every denomination

in Korean Buddhism. There were many schools during his time: Hae Dong Chong (Eastern School), Jeong Do Chong (Middle Path School), Bup Seong Chong (Dharma Nature, Emptiness School), Jeong To Chong (Pure Land School), Hau-Yen School (Avatamsaka School) and Bun Hwang Chong (Bun Hwang was his monastery) and Yul Chong (Vinaya sect, Precept School).[31] However, Wonhyo himself never established any school

If Wonhyo's theory of Reconciliation became widely known, the results in terms of achieving brotherhood and peace through mutual understanding among different religious sects could be quite significant. Wonhyo's theory of Reconciliation, with its emphasis on the one universal truth contained in all the various sectarian interpretations, could form the basis not only for a syncretic Buddhism and eventually even hold the possibility for a universal religious amalgamation, but also provide the rationale for a dialectic between religious groups and society.

Wonhyo (AD 617–686) was one of the most creative thinkers in the Korean Buddhist tradition. The time during which Wonhyo lived witnessed the unification of the Korean Peninsula under the Shilla Dynasty. This great monk lived through a year in Korean history that was characterized by social change and cultural amalgamation. It has been written that one cannot hope truly to understand traditional Korea without some understanding of Wonhyo, for he reflected his own time as well as moulded the path for future Korean philosophical development.[32]

Under the Shilla dynasty, various parts of Korea were unified and brought into sustained contact with each other for the first time and there occurred a synthesis from which Korean national identity arose.In a similar fashion. Wonhyo attempted to overcome Buddhist sectarian fighting and expound a theory of liberation free from dogmatic arguments.

Fundamental to all of Wonhyo's works is the idea of doctrinal reconciliation.[33] This syncretic tendency was most explicitly elaborated in his work entitled the *Treatise* on the *Reconciliation of all Disputes in Ten Aspects* (hereafter referred to as the theory of reconciliation). Frequently schoolars have cited this text as crucial in understanding the essence of Wonhyo's thought. His thinking formed the foundation of subsequent Korean Buddhist philosophy. The fact that Korea for many centuries was the only Buddhist country to have only one Buddhist sect, the non-denominational Chogye Order, is a lasting tirbute to the wide acceptance within Korea of Wonhyo's non-sectarian philosophy.[34]

Wonhyo wrote commentaries on the essence of the sūtras belonging to the individual sects. (Each school generally bases its doctrines on a particular sūtra). In these commentaries, he tried to capture the essential points of the lengthy, involved Sūtras in a simple form so that many people could understand the essential meaning of these texts. In addition, Wonhyo tried to show that the same essential Buddhist philosophy

formed the basis of each sūtra and that the only differences that existed between the various sūtras and, therefore, the school that embodied the philosophies contained in these sūtras was in human interpretation, not doctrine. Thus, Wonhyo reconciled all the so-called 'different' Buddhist doctrines and disputes, showing that there is only one essence of Buddhism.

Wonhyo stated in his *Taesung Kishinlonso* (The Commentary on the Treatise of Awakening Mahayana Faith):

> If one unfolds (the truth), the infinite and endless aspects become (various) doctrines, if one folds it up, the One Mind with two gates (of Suchness and Birth-and-Destruction) becomes the essence. Within the two gates, myriad aspects are embraced without confusion. Therefore, unfolding and folding are free and establishment and destruction are without obstruction: Unfolding is without confusion, and folding is without narrowness: establishment is free from impediment, and destruction is free from loss.[35]

According to Wonhyo's many commentaries and interpretations, we must find the Chong Yo; Chong represents the diversity in Buddhist truth by unfolding (deductive method) the peculiar characteristics of each literature or doctrine: yo represents the fundamental unity of truth by the logic of folding (inductive method). Fundamentally, Wonhyo's logic in the theory of Reconciliation is based on the philosophy of Chong Yo (doctrines and essence) and Kae-hap (unfolding and folding or sealing), which means he used both the inductive and deductive method without any prejudice. This methodology of logic has contributed very much to Buddhist doctrinal development.

Doctrine refers to the development of the one in the many, while essence refers to the unification of the many into one. When the truth unfolds, it is called doctrine: when sealed or folding, it is called essence. The unfolding aspect of the truth is also called the arising aspect of dharmas, while the sealing or folding aspect of the truth is called the ceasing aspect of dharmas.

The unique characteristics of Korean Buddhism can be traced to Wonhyo's philosophy. Wonhyo reconciled the contradictory assertions of the Hīnayāna and Mahāyāna schools and unified their varying philosophies in the theory of One Mind as the origin of all. One Mind is the same as the mind of pure self-nature, true suchness, Buddha Nature, Dharma Body, Tatāgatha store consciousness, Realm of Law and Dharma Nature, Dharmadhatu, the Realm of Reality. This means that mind is the ultimate reality of Buddhism.

Everything is created by mind alone. One Mind can also be described as the mind of non-discrimination and non-thought. As soon as one returns to One Mind, one no longer discriminates between sentient beings and Buddhas or between enlightenment and non-enlightenment,

but instead directly cognizes the world of emptiness and dependent origination.[36] In Buddhist philosophy, to return to One Mind and realize the Suchness of all phenomenal things is to obtain a true faith that 'I am Buddha'. As a result, Korean Buddhism had been called 'Unified Buddhism' as compared to the 'Original Buddhism' (before the emergence of sectarianism) of India and 'Sectarian Buddhism' of China.

Wonhyo's thought and way of practicing have influenced the Buddhism of China and Japan. Also,in Korean Buddhist history, Wonhyo's thought had a continuing influence and brought about the idea of National Teacher Dae Gak (Uicheon, AD 1055–1101) of 'Reconciliating meditation and doctrine'.[37] His influence is also reflected in the philosophy of National Teacher Bojo (Chiunl, AD 1158–1210) of 'returning to Dhyana (Zen) after studying doctrines.[38] Most Buddhist scholars agree that the three most crucial scholastic figures in Mahayana Buddhism in India, China and Korea are: Nagarjuna (AD 150–250) in India, who established the foundation of Mahāyāna Buddhism.[39] Chi-i (AD 538–597) in China, who unified the conflicting Three Vehicles in One Lotus schools[40] and Wonhyo in Korea, who harmonized all conflicting Buddhist sects and theories.

Wonhyo's philosophical dissimilarity with both Nāgārjuna and Chi-i led to the foundation of Hoe-Tong Bulkyo (Buddhism of Unity and Interpenetration), which is based on the theory of the Reconciliation of all disputes, which attempts to manifest the main philosophical teaching of Buddha, united into one truth. This is the fundamental concept of Wonhyo's thought and philosophy.[41] Accordingly, Wonhyo's posthumous title was Hwa Jaeng Guksa (National Teacher who harmonized all disputes).[42]

V. THE SIGNIFICANCE OF WONHYO'S PHILOSOPHY ON MODERN VALUES

Wonhyo's philosophy of Reconciliation was eventually successful in uniting all the Buddhist schools in Korea into one school—the Chogye Order—in the fourteenth century during the Korean Yi dynasty. This unified Chogye Order of Buddhism still exists in Korea today.

During the seven hundred years following his death, Wonhyo's philosophy contined to influence Buddhist leaders in Korea. During these years there were three patriarchs who were particularly influenced by Wonhyo's thoughts on Reconciliation. These leaders were Koryo Dynasty National Teacher Uicheon (AD 1055–1101), Chinul (AD 1158–1210) and Seo-San (AD 1520–1604). As a result of Wonhyo's teachings on Reconciliation, the character of Korean Buddhism is uniquely different from that of China or Japan. According to Korean Buddhism, written by Nam-Son Choi, Korean Buddhism is characterized as 'Tong Bulkyo (unified, synthetic, and/or universalistic Buddhism), 'Buddhism of total interpenetration' or 'concluding Buddhism' in contrast with the 'introductory Buddhism' of India and the 'particularistic or sectarian Buddhism' of China.[43]

What is the value and meaning of the logic of Wonhyo's theory of Reconciliation for the modern world and our present age?

The basis of the theory of Reconciliation is the principle of unification and non-discrimination, which could be the basis for creating a world of equality and peace. The modern world faces a tense situation resulting from the sharp confrontation of different religious, philosophical, economic and political systems. If human beings hope to establish world peace, as reflected in an ecumenical movement transcending all sectarian beliefs and philosophical differences, it would be helpful to study Wonhyo's theory of Reconciliation.

NOTES AND REFERENCES

1. Mrs Rhys Davids, *Gradual Sayings*, Vol. 1, p. viii.
2. The Stcherbatsky, *The Central Conception of Buddhism*, Motilal Banarsidass, Delhi, p. 73.
3. P.T. Raju, *Idealistic Thought of India* p. 282.
4. Ibid., p. 285.
5. C. Sharma, *A Critical Survey of Indian Philosophy*, p. 70.
6. *Majjhima Nikāya*, 63.
7. *Samyutha Nikāya*, iii.
8. *Mādhyamika-Kārika*, opening verse.
9. *Majjhima Nikāya*, ii. 32; S. Radhakrishanan, *Indian Philosophy*, Vol. 1, p. 371.
10. The *Dhammapada*, 393.
11. *Mahāparinirvāna Sūtra*, Vol. VI, p. 1. XV, 6.
12. Dae Han Pulkyo Chong Yonmeng (ed.), *The Teaching of Buddha* (Seoul; Hanjin Publishing Company, 1979), pp. 49–55.
13. P.T. Raju, *Idealistic Thought of India*, p. 279.
14. P.T. Raju, *The Philosophical Traditions of India*, pp. 122–123.
15. P.T. Raju, *op. cit.* p. 235.
16. Ibid., pp. 235–36.
17. Suzuki (trans.), *The Awakening of Faith*, New York, 1967, p. 43.
18. Yamakami Sogen, *Systems of Buddhist Thought*, Calcutta University Press, Calcutta, 1912, p. 253–251.
19. Suzuki, op. cit., p. 57.
20. Ibid., p. 58.
21. Ibid., pp. 58–59.
22. Ibid., p. 59.
23. P.T. Raju, *Idealistic Thought of India*, p. 228.
24. *Ibid.*, p. 230.
25. G. Tucci, *Prajñāpāramitās*, Vol. 1, p. 218.
26. Ibid., p. 372.
27. Ibid., p. 272.
28. English translations by Takakusu, p. 148 (S.B.R., Vol. XLIX).
29. Young'tae Kim, *Han' guk Bulkyosa Kae Seol*, Introduction to the History of Korean Buddhism, Kyong Seo Won, Seoul, 1986, p. 23.
30. Cheol-u Mok. (ed.), *Wonhyo Jeonjip* (Wonhyo's Complete Collection), Buddhist Studies Association, Donagguk Yokkyong Won, Dongguk University, Seoul, 1973,

p. 303; and Sung-Bae Park, '*Wonhyo's Commentaries on the Awakening of Faith in Mahāyāna*' (doctoral dissertation, University of California, Berkeley, 1979, p. 74.

31. Myong-gi Cho, *Shilla Bulkyo ui Inyon Gwa Yoksa*, The Ideologies and History of Shilla Buddhism, Seoul: Shin Taeyang Sa, Seoul, 1962, p. 248.
32. Myong-gi Cho, *Shilla Bulkyo ui Inyon Gwa Yoksa*, The Ideologies and History of Shilla Buddhism., p. 238: and Chong-Hong Park, *Han'guk Sasangsa*, The History of Korean Thought, Somundang, Seoul, 1972, p. 25.
33. Chong-Ik Lee, *Wonhyo ui Kunbon Sasang*, Wonhyo's Fundamental Philosophy, Tongbang Sasang Yon'guwon, Seoul, 1977, pp. 10, 68.
34. Young't'ae Kim, *Han'guk Bulkwosa Kaw Seol*, Introduction to the History of Korean Buddhism, Kyong-Seo Won, Seoul, 1986, p. 77.
35. Ibid., p. 213: Hee Sung Keel, *Chinul, the Founder of Korean Zen Tradition* (doctoral dissertation, Harvard University, 1977, p. 39; and Chong-Hong Park, *Han'guk Sasangsa*, The History of Korean Thought, p.87.
36. Dependent origination is the law of causation. According to this doctrine, no phenomenon arises without a cause. Things and acts are not isolated but are connected with other things and acts in the universe. One acts influences on another and is in turn influenced by others.
37. Chong-Hong Park, *Han'guk Sasangsa* (The History of Korean Thought), pp. 148–162.
38. Chong-Ik Lee, *Dong Bang Sasang Nanchong* (The Collection of Eastern Thought), Boryonkak, Seoul, 1975, pp. 264–272.
39. Kenneth K. Inada, *Nāgārjuna—A Translation of His Mūlamadhyamakakārikā with an Introductory Essay* (Tokyo: The Hokuseido Press, 1970), p.3.
40. Chong-Ik Lee, *Bulkyo Sasang Dae Kwan* (The collection of Buddhist Thought), Boryonkak, Seoul, 1973, pp. 256–259. Three vehicles (*Triyama*), or conveyances which carry living beings across *saṁsāra*, or mortality (births and deaths) to the shores of *nirvāṇa*.
41. Myong-gi Cho, *Shilla Bulkyo ui Inyom Gwa Yoksa*, pp. 221–225.
42. Ibid., p.92; and Chong-Ik Lee, *Dang Bang Sasang Nonchong* (The collection of Eastern Thought), p. 220.
43. *Yuktang, Choi Nam Son Jeonjip* (Complete Collection of Choi Nam Son), Seoul, 1973, pp. 547–72.

Theory of *Vākya* or Proposition

V. SHEKHAWAT
Department of Philosophy, University of Rajasthan, Jaipur

1

Saṁvāda Gaṇita or *Pratīkanvīkṣikī* differs from *Gaṇita* or *pratīka bhāṣā* essentially in this that it seeks to symbolize the argument-units themselves which cannot, however, be symbolized adequately unless the *pramāṇas* are symbolized. Now, *saṁvāda* restricts itself only to genuine knowledge by adhering to the principle of *pramāṇa*. Genuine knowledge can however be erring or unclear or ambiguous so that validity of *saṁvāda* is preserved by adhering to the principle of *siddhi* whereupon true premises, whether affirmative or negative, yield true conclusion only. *Saṁvāda*, then, consists in drawing fourfold inferences, namely *utsargakaraṇa, anumānakaraṇa, upamānakaraṇa*, and *yuktikaraṇa* each of which is itself of manifold kinds.* The basic unit of which an *utsarga/anugamana* or *anumāna* or *upamāna* or *yukti* is itself composed is the *proposition* which is true or false, that is presenting genuine knowledge. In order thus to symbolize the *pramāṇas* and the *yukti* so as to present *siddhi* as derivational structure, we shall have to seek the symbolization of proposition units themselves. The fundamental questions then are: How many kinds of proposition units are there? And how to symbolize them?

A question more fundamental than the above, however, is: How is a proposition-unit generated? One may say that a proposition-unit is generated by *vyākaraṇa* or *vyākṛti* of *prākṛta vāgbhāṣā*. This is however only partially true for *saṁvāda gaṇita* recognizes that although a proposition may be *vyākṛta* in so far as norms of *Vyākaraṇa Śāstra* are concerned, it may be 'avyākṛta' from the perspective of *saṁvāda gaṇita*. For example, '*Puruṣa nitya hai*' is *vyakṛta* in so far as *Vyākaraṇa Śāstra* is concerned but not so in *saṁvāda gaṇita* for it is purported to be a universal proposition of which the *vyākṛta saṁvāda* form will be '*Puruṣa nitya hotā hai*'. Because of logical ambiguities that *Vyākaraṇa* permits, we hold that a proposition-unit is generated from *statements* by means of *utsargakaraṇa*. *Utsargakaraṇa* is therefore not merely a schema/principle for validly inferring universal propositions from particular ones but also a schema/principle for *generating siddha* particular proposition-units from statements. A statement

* See V. Shekhawat, 'Problems of Formalization in *Saṁvāda Śāstra*', *JICPR*, XIII (2), 1996.

is a *natural expression* which *presents roughly* an *instance*, or, it is a *vyākṛta* expression which presents an instance or situation in accordance with norms of *Vyākaraṇa*. As different from this, a proposition *depicts* or *composes clearly* a fact whether fully or partially. We may say that a statement or *kathan* guarantees only genuine knowledge in so far as it is *selected* by *pramāṇa* whereas a proposition or *vākya* guarantees *errorless* genuine knowledge—whether affirmative or negative. It is by *utsarga/anugamana* that we *generate* or reach a particular or universal proposition by means of several similar statements. Therefore statements as such cannot be allowed to enter any knowledge-system unless these are established/proved as propositions. Clearly, *saṁvāda vyākṛti* is different from *vāgbhāṣā vyākṛti* in so far as it demands the generation of propositions not statements; the former is over and above the latter, so to say.

When we thus generate propositions from statements by means of *utsargakaraṇa*, can a singular statement be admitted as *apavāda*? Indeed a singular statement can make one inquisitive as to the *possibility* of *apavāda* and thus spur one for exploration/inquiry in the direction pointed out by the statement and this provides it a limited *value* in knowledge systematisation for a single instance *can* often turn out to be a massive challenge to some accepted system of knowledge. Yet we must ensure that only sufficiently numerous *apavāda*-statements generate an *apavāda*-proposition whether singular-*apavāda* or multiple-*apavāda*.

Clearly, the most difficult task in *Saṁvāda Śāstra* is, firstly, to *generate* propositions or *vākyas*, and, secondly, to ensure that these are *true*. In *Saṁvāda Śāstra*, we are not interested in investigating under what conditions facts *occur*, rather we are interested in investigating under what conditions the *knowledge* of facts is *true*. Strictly speaking, this is not the business of *logic* as contemporaneously understood but this has great significance in *Saṁvāda Śāstra* as it provides us *models* for adequate *yuktikaraṇa* which is a strictly logical enterprise. Epistemology then forms the foundation of Logic; it is the *Jñāna-anvīkṣā* of *saṁvāda*. that is, our conception of a proposition, methods of generating propositions, contradictory of a proposition, universality of propositions, their *apavāda* forms etc. [Indian thinkers might have failed to make a headway towards formalization of *saṁvāda* because they remained too much preoccupied with the first project of epistemological grounds/foundations,] Strictly speaking thus, when we attempt to formalize *pramāṇakaraṇa*, we are attempting the formalization of the epistemological component of *Saṁvāda Śāstra* not its logical component, a point often missed by contemporary formalisers of the *Śāstra*.

II

Now, propositions can be elementary, simple or compound (*rāśi*). Elementary proposition is *natural* and its property of being true or false

is also natural in so far as it gets selected by the natural process of *pramāṇa* in the *loka*. '*Yeh phool hai*', '*Yeh neem hai*', '*yeh peḍa hai*', etc., are elementary propositions, '*Yeh neem kā peḍa hai*' is a simple proposition. Elementary proposition is elementary in the sense that it does not depict any *relational* structure, it is uncomposed being natural. A simple proposition on the other hand depicts *at least one* simplest relational structure. '*Parvata para agni hai kyonki vahān dhuān hai jaise rasoighara men hotā hai*' is the example of a compound/complex proposition or a proposition-*rāśi*. It is compounded of elementary propositions: (*yah parvat hai*) (*yahān agni hai*) (*yahān dhuān hai*) (*rasoighara hotā hai*) (*dhuān hotā hai*) and connectives '*kyonki*', '*jaise*' as also of relations '*para*', and '*men*'. '*Yeh dhuan hai*' and '*dhuān hotā hai*' are both to be treated as elementary propositions, the former a *singular* elementary proposition, and the latter a partially universal elementary proposition. The latter is only partially universal because in it '*hotā hai*' occurs without suggestion of thorough universality, it merely suggests a *temporal recurrence* or mere existence without any commitment whatsoever to universality as generally understood.

Elementary propositions generally occur as a rule only in *utsargakaraṇa*. An elementary proposition is so because it depicts a simple situation and a complex proposition is so because it depicts a complex/compound situation. An elementary proposition is *somewhat* deficient of depiction/composition and a complex proposition *completely* depicts/composes a situation. We may say a simple proposition is *somewhat complete* depiction of a situation, it is a *mādhyam*-proposition. A simple proposition depicts *only* by means of relations(s) whereas a proposition-*rāśi* depicts by means of *relations* as well as *connectives*. The simplest of simple propositions is that which has just one relation and just one *udeśya* and just one *vidheya*. When we analyse a proposition *rāśi* into elementary propositions, it does not mean the latter just *join-up* to make the former a *rāśi*,—being a composition unit *rāśi* is an organic *whole* composed of elementary *parts*. Clearly, concepts or *term* (*uddeśya, vidheya*), *relation* and *connective* are fundamental to the understanding of a relational structure,—that is understanding it *logically*. If one asks 'what is the structure of an elementary proposition?' the reply could be that it is a question beyond the interest of *logic*, it rather lies in the purview of *grammar*. And we are not interested in grammatical structure of propositions, we are interested only in their logical structure. It may well be said that elementary propositions are *logical* without having any logical *structure*,—these have only grammatical structure.

Certain depiction/composition of a situation may have correct *grammatical form* only such as '*yah peḍa hai*' '*yeh neem kā peḍa hai*', '*usa parvata para dhuān hai*'. Another depiction may have both correct grammatical form as well as correct *logical form* such as '*neem kā peḍa hai*', '*parvata para dhuān hai*'. The latter will be called *correct* simple propositions or *right* simple propositions because these are simple propositions that are *grammatically as well as* logically right/correct. We may say that it is the

pramāṇakaraṇa that *transforms* simple statements (that is depictions or, better, expressions, having correct grammatical form only) into correct simple propositions,—particular or universal, the exception/*apavāda* being the elementary propositions only which pass untransformed. In *yuktikaraṇa*, then, we allow, as far as possible, only correct simple/compound propositions whether particular or universal.

Clearly, an elementary proposition is also an elementary statement,—the two are in distinguishable which therefore is the proof of elementariness of depiction/expression. It follows that in so far as analysis of elementary statements/propositions is concerned, logic coalesces with grammar. And as far as epistemology is concerned, elementary statements/propositions are *natural* and given to all humans without exception. Thus, in elementary situations such as '*yaha neem hai*', '*yahān neem hai*', we cannot logically distinguish between '*yah*', '*yahān*', etc., though grammatically these express differently: '*yeh*' or '*yah*' refers to *neem* while '*yahān*' refers to the *place* that *neem* occupies, etc. The variations, so to say, collapse in the same simple situation of 'this tree being here' to be depicted by the elementary statement/proposition in which the only term that matters is *neem*.

Unlike elementary propositions, simple/compound propositions can be classified further as one-relation, two-relation etc. propositions or one-relation-one-connective, two-relation-one-connective, etc. propositions. Thus, '*neem kā peḍa hai*', '*neem kā peḍa harā hai*', '*harā neem kā peḍa jhoola rahā hai*', '*neem kā peḍa havā se jhoola rahā hai*' etc. And similarly compound propositions such as '*parvata para dhuān hai kyonki agni hai*', '*parvat par dhuān hai kyonki agni hai jaise rasoighara men dhuān hotā hai*', etc. There are thus numerous varieties of propositions. Propositions that are most suitable for symbolization are one-relation simple propositions and two-relation-one-connective compound proposition-*rāśis*.

Now, a proposition is *true* by virtue of the case being *actually* so—a true proposition is the *artha*/fact *as said*—but is false by virtue of *actual* absence of *uddeśya* or *vidheya* as well as by virtue of their replacement by some other *uddeśya* or *vidheya*, not merely by virtue of the case *not being actually so*. However, we ordinarily require only this condition of falsity that it is the actual *absence* of a feature *alone* that makes a proposition asserting its presence false, not the presence of different feature which in invariably there for if that requirement were added, the presence of a feature different from the said different feature may or may not make it false. Therefore, just as specific presence makes a proposition true, a specific absence makes it false. This conception of falsity has the advantage of *symmetry* with the conception of truth. This however, makes it abundantly clear that an affirmative proposition depicts the entire situation/fact whereas its negative counterpart depicts only half the situation/fact as it always *points to some positive feature* that has been left undepicted/unmentioned. In this sense only positive/affirmative propositions have

strict claim to *truth*. For these depict the actual relation between two *presences*. Negative propositions are only half-claimants to truth for these only inform about the relation between a presence and an absence but not about the other relation that actually obtains between the said presence and some other (unsaid) presence. It follows that to treat affirmative proposition and its negative as alternating *states* (of truth and falsity) is somewhat misleading. Thus, for example, '*neem harā nahīn hai*' is said to be true by virtue of relation of '*neem*' with absence of '*harā*' but that is only partial truth, the full-blown truth comes out when it is said that '*Neem pīlā hai*'. Strictly speaking, then, if '*neem harā hai*' is true, its 'negative' is '*neem pīla hai*' as it depicts a full-blown falsity of the first situation. Relation of a presence and an absence thus only *implies* a relation of *that* presence *with some* presence at least in simple propositions where certain specific relations obtain. Similarly, in some other relations occurring in simple propositions, relation between an absence (such as a *sādhya* or *hetu*) and a presence *always* simplies relation of *some* presence with *that* presence. It follows that in so far as simple propositions are concerned, we should allow minimum number of negative propositions in any system of knowledge for these always leave one in *suspense* as to what the situation *actually* is and are thus ambiguous/suspense statements. Thus, even when we present a proposition counter to any affirmative, we try to make it counter-affirmative rather than negative, for example, the counter-affirmative of '*neem harā hai*' is '*neem aharā hai*'.

The situation, however, is not the same for some elementary negative propositions. These have the remarkable property of being *categorically* negative as these clearly express the situation of absence without leaving one in any suspense whatsoever. Thus, for example, '*yahān neem nahīn hai*' categorically expresses the absence of *neem* and also makes the *logical distinction* from '*yeh neem nahīn hai*' clear for the latter elementary proposition is not categorical as it leaves one in suspense as to what it is if not *neem*? Categorical negative elementary propositions therefore have special logical value and are often interpreted as affirmative propositions presenting specific absence.

III

Consider now the following propositions:

(i) *Sabhī parvaton par dhuān hotā hai.*
(ii) *Sabhī parvaton par dhuān hai.*
(iii) *Parvaton par dhuān hota hai.*
(iv) *Kucha parvaton par dhuān hotā hai/rahatā hai.*
(v) *Kabhī-kabhī parvaton par dhuān hotā hai.*
(vi) *Usa parvat par dhuān hotāhai/rahatā hai.*
(vii) *Kucha parvaton par dhuān hai.*

Of the above, proposition (i) is a maximally universal proposition for it asserts that on all the mountains for all *times* there is smoke. In other words, the proposition asserts both *numeral* as well as *temporal* universality.[1] Not so with proposition (ii) and proposition (iv) as the former asserts only numeral universality at present while the latter asserts only temporal universality on a single mountain,—this latter one is apparently a singular proposition yet hides in it *temporal universality* in a single instance. Proposition (iii) may be said to be a just universal proposition for although it does not make claims of maximal universality, yet it asserts that on *most* mountains, by and large, regularly there is smoke. Proposition (iv) is, strictly speaking, an *apavāda*-proposition for it merely does not assert that there is often smoke on *some* mountains but also suggests that there are *even more* mountains on which there is no smoke. The proposition asserts numero-temporal *apavāda* of proposition that would be maximally universal asserting variant situation. As different from this, proposition (v) asserts only temporal-*apavāda* not committing itself numerally—that is *apavāda* may be true for all or some mountains. And proposition (vii) asserts only numeral-*apavāda* of the present. This analysis makes clear the distinction between maximal, numeral and temporal universality on the one hand and numero-temporal, numeral and temporal *apavāda* or exceptionality on the other hand. Proposition (vi) provides a singular-*apavāda* or local universality by committing to temporal oftenness or recurrence and is different from '*kabhī-kabhī usa parvat par dhuān hotā hai*' which ecpresses strictly singular temporal-*apavāda*. What is important to note is that *apavāda* itself often involves temporal-universality and requires not *at least one* instance—numerally or temporally—but *more than one* instance in order to make sense and be distinguishable from particular propositions such as '*Usa parvat par dhuān hai*' which merely asserts a situation as it obtains at present. That is to say, the cosmos is such that *particular* instances acquire the status of numeral-temporal-*apavāda* of above universality when observed more than *once* of more than *one* by more than *a person*.

It was shown earlier that a negative particular proposition falsifies its affirmative counterpart although it asserts only partial truth regarding what the situation is like and only an affirmative variant proposition with same *uddeśya* but different *vidheya strictly* falsifies the original affirmative proposition. The situation regarding affirmative universal propositions is however more complex not only because of diverse universalities but also because of restricted universalities of the *apavāda* as variants. There can be several kinds of *falsifying situations* for an affirmative universal proposition, such as:

(i) a negative particular proposition,

[1] This is the generally accepted depiction. A clearer depiction will be '*sadāsarvadā parvaton par dhuān hotā hai*'.

(ii) an affirmative particular proposition,
(iii) a negative *apavāda* proposition,
(iv) an affirmative *apavāda* proposition,
(v) a negative of the universal proposition,
(vi) an affirmative universal proposition,
(vii) an affirmative counter-proposition.

A thorough analysis of conditions of falsification of an affirmative universal proposition thus calls for analysis of *terms* of proposition and *relations* between these terms.

IV

Now, the terms of a proposition can be particulars or universals and behave so depending on the structure of the proposition. Thus, in the proposition '*neem harā hai*', both *neem* and *harā* are particulars but in '*neem harā hotā hai*', both the terms are universals. Moreover, in '*sabhī neem hare hote hain*', the terms *neem* and *hare* are interpreted as classes. Numerous kinds of relations obtain between terms of propositions but four are generally recognized as fundamental, namely *samavāya, tādātmya, saṁyoga* and *sādṛśya*. (The more the kinds of relations considered in a system of logic, the more comprehensive it would be). Thus, in '*neem harā hai*', the relation between the terms is *samavāya* relation whereas in '*sabhī neem hare hote hain*' it is *tādātmya* relation. Similarly in the proposition '*parvata par dhuān hai*', the relation between the terms is *saṁyoga* relation whereas in the proposition '*yah gāya jaisā prāṇī hai*', the relation is that of *sādrśya*. Now, a negative particular proposition is invariably non-categorical when the relation between its terms is a *necessary* relation, that is either *samavāya* or *tādātmya*. Thus, for example '*parvata par dhuān nahīn hai*' is a *categorical* negative particular proposition but not the negative particular proposition '*neem harā nahīn hai*'. It is clear that at least for propositions of which terms have non-necessary relations, the corresponding *negative* propositions are decisive *falsifiers*. And this can be accepted as a rule for universal propositions also provided we settle the question what *quantity* will suffice for falsifying it. Now, a particular affirmative or negative proposition cannot be a decisive falsifier of an affirmative universal proposition because the falsifying instance may be due to some peculiarity of the situation. Neither can the affirmative or negative *apavāda*-proposition be a decisive falsifier for it can be the peculiarity of some *class* of particulars which may be formulated as a restricted universality as addendum to the parent affirmative universal proposition. Thus, for example if '*sabhī vṛkṣa hare hote hain*' is an affirmative universal proposition, its particular negative falsifier is '*vṛkṣa harā nahīn hai.*' and particular affirmative falsifier is '*vṛkṣa pīlā hai*' both being not decisive; its negative *apavāda* falsifier is '*Kucha vṛkṣa hare nahīn hote*' and affirmative *apavāda*

falsifier is '*Kucha vṛkṣa pile hote hain*', both again not being decisive as the latter can be formulated as a restricted universality '*amaltāsa pīle hote hain*', the term *amaltāsa* being included in the term *vṛkṣa*. Thus the original parent universal proposition can still be maintained as *true* by reformulating it as '*sabhī vṛkśa hare hote hain parantu amaltāsa pīlā hotā hai*'. It follows that only a negative universal proposition or an affirmative universal proposition can be the *decisive* falsifier of the parent universal proposition, the former in cases where the relation between terms of the parent proposition is non-necessary and latter in cases when it is necessary. Thus, the *decisive falsifier* of '*sabhī vṛkṣa hare hote hain*' is '*sabhī vṛkṣa angūrī hote hain*' and of, for instance, '*sabhī parvaton para dhuān hotā hai*' is '*kisībhī parvata para dhuān nahīn hotā*'. That is if '*sabhī vṛkśa hare hote hain*' is *true* then '*sabhī vṛkśa angūrī hote hain*' *must* be *false*; and if '*sabhī parvaton par dhuān hotā hai*' is *true* then '*kisībhī parvata para dhuān nahīn hotā*' *must* be *false*; all the rest of falsifiers may yet be true when the parent proposition is true.

A negative universal proposition depicts an *empty class* when the relation between its terms is not necessary relation, but not otherwise. Thus '*kisībhī parvata par dhuān nahīn hotā*' categorically depicts *absence* of smoke on all mountains. And it is possible to have a doubly negative universal proposition such as '*na parvata hotā hai na dhuān hotā hai*' which depicts two absences at once, namely, of mountain as well as smoke; this sort of proposition differs from '*Esā koibhī parvata nahīn hotā jisa para dhuān nahīn hotā*' which is precisely an affirmative by double negation. When one or two absences/empty-classes are depicted categorically, what sort of relation holds between the terms? In the case of a single absence and a particular/universal, we presume that the same relation—in the above case *saṁyoga* relation—would hold as between two presences. However, when two absences are categorically depicted, we shall have to conceive of a special relation between absences.

If we write *dhuā nahīn* as *na-dhuān*, then this *negative term* differs from the complementary term *adhuān* for while *na-dhuān* depicts *absence* of smoke, *a-dhuān* depicts a class from which smoke is *excluded*, a distinction that is not often noticed in received *saṁvāda śāstra*. Thus, to say that '*Puruṣa nitya nahīn hotā*' is not the same as saying '*Puruṣa anitya hotā hai*' for if we interpret these propositions as depicting relations between classes, the former depicts a relation between the class of *puruṣas* and absence of class of *nitya* objects while the latter depicts a relation between class of *puruṣas* and class of objects that excludes *nitya* objects. Complementary terms can be as ambiguous as negative propositions of which terms have necessary relation for these also leave us in suspense as to what objects belong to the class if not the excluded one? In a sense, propositions such as '*Sabhī neem alatā hote hain*' depict a sort of 'absurdity' for if *alatā* happened to be a class that excludes *latā*, then *neem* would seem to belong to all classes except *latā*'.

What we have undertaken above are only preliminary considerations regarding universal propositions and their relation to particular falsifiers and *apavāda*-falsifiers and it is obvious that a detailed theory of various kinds of universal propositions and *apavāda* propositions is itself called for.

V

While attempting symbolization of propositions, we are to remember that the terms in each of the propositions are related in a *definite order*. Thus '*neem harā hai*' means *harā* inheres *in neem* not vice versa and '*parvata par dhuān hai*' means *dhuān* is *on* the *parvata* and not vice versa. Some propositions having *samavāya* relation between terms are '*kaṇāda khātā hai*', '*neem kā peḍa hai*', '*kachue kī gardan hai*', '*Bīja se ankuraṇa hai*', '*agni se dhuān hai*', etc. Some propositions with *saṁyoga* relation are '*Puruṣa ko bandha hotā hai*', '*Prthivi para jala hai*', '*Ātmā men man hotā hai*', '*Kāla men ātmā hotā hai*', etc. Examples of propositions with *tadātmya* relation are: '*Sinśipsā vṛkśa hote hain*', '*Paśu prāṇī hote hain*', '*Puruṣa nitya hote hain*', etc. Examples of *sādṛaśya* relation are: '*Gāya sadṛaśa prāṇī gau hota hai*', '*Nābhika būnda-sadṛaśa hote hain*', etc. We shall employ x, y, z to represent variables and p, q, r, u, v, w, k, etc., to symbolize terms. Thus a singular proposition (or elementary proposition) such as '*yeh neem hai*' can be symbolized straight off as n_x. *Tādātmya* relation will be symbolized as in contemporary logic so that '*sabhī sinśipā vṛkṣa hote hain*' can be symbolized as (x) (y) $<s \supset v>_{xy}$. *Samyoga* relation will be symbolized by the symbol $\cup$ (bow) so that '*parvata para dhuān hai*' can be symbolized as $<p \cup d>$. *Saṁavāya* relation will be symbolized by Ξ (joint bows) so that '*neem harā hai*', etc. will be symbolized as $<n \Xi h>$, $<K \Xi k>$, $<n \Xi p>$, etc. Similarly, *sādṛśya* proposition will be symbolized as $(g_x \subset p_x)$ etc. In all these propositions, the first term/symbol presents *uddeśya* and the second symbol/term presents *vidheya*. Since the terms themselves do not possess the property of being true or false, only propositions do, the truth/falsity of propositions is not determined in any way by the terms. However, the terms serve to make explicit the *relational structure* of the proposition, that is, these present to us *how* these stand *related* to each other in a certain *order*—an advantage often absent in modern symbolic logic. The one and only condition of *truth* of the proposition is that the two terms *actually* stand related the way these are depicted/composed. However, as regards the conditions of *falsity* of propositions, these are three: (i) the *uddeśya* is missing actually, (ii) the *vidheya* is missing actually and (iii) both *uddeśya* and *vidheya* are missing actually. In *tādātmya* relation, however, there will be further conditions such as inclusion as asserted does not obtain and similarly in case of *sādṛśya* relation. Since we want only true propositions to enter/occur in our *saṁvāda*, we only inquire about the *truth condition* of propositions (by means of *parmāṇakaraṇa*) and only when some false

proposition enters by mistake in our *saṁvāda* do we inquire about its falsity conditions so as to expel it decisively.

In case of universal and *apavāda* propositions, we will symbolize numeral universality by (x) and temporal universality by (y), so that the proposition '*sabhī parvaton para dhuan hota hai*' will be symbolized as (x) (y) $<p \cup d>_{xy}$. Just universality will be symbolized by U* so that '*parvaton par dhuān hotā hai*' will be symbolized as U* $<p \cup d>$. '*Sabhī parvaton par dhuān hai*' can be symbolized as (x) $<p \cup d>_x$ and '*Isa parvat par dhuān hotā hai*' as (y_1) $<p \cup d>_{y1}$, '1' indicating singularity, '*Kucha paarvaton par dhuān hotā hai*' will be symbolized as $\exists_x \exists_y <p \cup d>_{xy}$, and '*kucha parvaton par dhuān hai*' as $\exists_x (p \cup d)_x$ etc, etc.

Sādṛśya relation is symbolized as $\subset$, so that the proposition '*nābhika bŭnda-jaisā hotā hai*' will be symbolized as (y) $<n \subset b>_y$ and '*Gāya-sadṛśa prāṇī hai*' as $<g \subset p>$.

It is clear that according to *saṁvāda śāstra*, when we undertake to construct a *siddhāntatantra*, we ought to follow the policy of, firstly, selecting, by means of *pramāṇa* criterion, genuine knowledge alone and, secondly, allow only *true* propositions to enter the *siddhāntatantra*. Moreover, these propositions—whether universal or particular—ought to be affirmative and/or categorically negative as far as possible.

VI

Consider now a proposition-*rāśi* such as this.: '*sabhī vyāpārī imāndāra hote hain parantu kucha vyāpārī kama tolte hain aur kucha vyāpārī milāvaṭa karate hain tathā kucheka māla dabāte hain*'. The question here arises: just how many *apavādas* can a universal proposition sustain? Or, how long can a universal proposition affirming a 'law' hold water or withstand in the face of *apavādas*? Certainly, an affirmation can hardly be said to be universal if it admits of too numerous exceptions. Thus, in our effort at theorization or *siddhāntatantrakaraṇa*, we follow the strategy of seeking, as far as possible, universal affirmations without any *apavāda* and when some *apavāda* appears, we get alerted about the universalization and moreover, as further *apavāda* appears the universalization is *suspect* and we start inquiring about the *deeper causal explanation* of the situation so as to obtain a new universal affirmation without *apavāda*. It is by virtue of this inbuilt demand of theorization that we are compelled to investigate into deeper causes of the world experienced as such.

And disallowing non-categorical negative propositions also serves this purpose of seeking deeper and deeper causes for as non-categorical negative propositions leave us in suspense as to what the thing is like if not so, we seek to resolve the suspense by seeking the affirmative. Similarly for counter-affirmative propositions. Now, *doubt* and *suspense* differ in this respect that while doubt is resolved merely by achieved *rejection* of a positive or negative alternative, suspense is resolved *only* by

making explicit the *unknown*. Thus in $p_x \vee \sim p_x$ if p_x is rejected logically/by means of *pramāṇa*, the doubt is resolved yet the suspense remains as to if the thing is not so what it is? Of course, in situations where such knowledge is not relevant, the suspense may never arise, for example one may not be inquisitive as to what there is on the hill if the smoke is not there, which is the reason for such propositions being categorical as *saṁyoga* relation does not invoke the presence of something else on the hill if there is no smoke, for it is likely that there is *nothing else* on the hill. Doubt, not suspense, withholds an erring mind from ascertaining a claim and *siddhi* eventually dispels error, therefore doubt,—and one can ascertain a claim of truth as well as falsity in the sense of non-categorical depiction.

Now, this exclusion by disjunction (as above) differs from exclusion by terms. In *tādātmya* relation, for example, complementary terms have one meaning (as being interpreted as classes) whereas in *saṁyoga* and *saṁavāya* relations there occurs a meaning variation/shift. Thus '*sabhī gulāba avṛkṣa hote hain*' is generally interpreted to mean that all *gulābas* belong to the class of things that the class of *vṛkṣa* excludes. But such cannot be the interpretation of '*sabhī neem ahare hote hain*' or of '*sabhi jalāśayon men adhuān hotā hai*'. If the class interpretation is adopted for these, the former says that objects excluded by the class *harā* inhere in *neem* and the latter that objects excluded by the class *dhuān* are in the lake, which makes clear that the interpretation is absurd. It may however be said that '*sabhī neem ahare hote hain*' claims only that the property '*hare*' does not inhere in *neem* whatever other property or properties may inhere in it, and this claim is sufficient to falsify the proposition '*sabhī neem hare hote hain*'. This is a perfectly legitimate claim, yet the claim does not resolve the problem of suspense as to if the *neems* are not *ahare* what property/properties they have? In other words, we have no right to utter *only* counter-affirmative propositions unless, at least subsequently, we *also* utter affirmative propositions so as to dispense with the suspense. In exclusion by terms, *harā-nahīn* or *na-harā* excludes only *harā* of the particular *uddeśya* but *aharā* excludes *all* the greens and, moreover, the former *refers* only to exclude *harā* whereas the latter refers to all those objects that are excluded by *harā*. Thus '*yahān ghaṭa nahīn hai*' provides knowledge of absence of particular *ghaṭa* whereas '*yahān aghaṭa hai*' provides knowledge of presence of a whole from which *ghaṭa* is excluded, that is of 'all that which is not *ghaṭa* and is *actually* present here'. Both *na-ghaṭa* and *aghaṭa* are conceived as properties/features of *field of perception*, the former informs of a *positive* feature namely absence of expected *ghaṭa* whereas the latter *nonexplicitly* informs of presence of things other than *ghaṭa*. Moreover, the latter presupposes a definite field of perception which, if not accessible, would disallow their presentation but such is not the case with the former which presupposes only the access to expected *ghaṭa*.

The exclusion by disjunction, however, differs from exclusion by

terms as shown above. Thus, for example, when one is in doubt as to whether '*yah muskurāyegā yā nahīn muskurāyegā*' or '*yah sinśipā hai yā nahīn hai*' or '*yah saṛpa hai yā rassī hai*', one entertains different sorts of situations. All these situations definitely do not presuppose any field of perception over and above the particular object of perception. Moreover, these do not provide any *definite* genuine-knowledge either, for their exclusion awaits further inquiry of occurrence which may never take place. It only suggests an *expected* exclusion therefore an *expected* availability of /access to proved genuine-knowledge. Strictly speaking therefore, the exclusion by disjunction provides only statements or *kathana* not propositions or *vākya*. It then belongs to an entirely different class of knowledge, namely of genuine-knowledge that *awaits siddhi* or occurrence. It may be noted that absence of *expected* occurrence and presence of occurrence present the same kind of alternatives for doubt as do absence of knowledge and presence of knowledge (such as in case of *sinśipā*). The example '*parvata para dhuān hai yā nahīn hai*' also belongs to this last variety of exclusions. The situation in case of '*yah sarpa hai yā rassī hai*' is however more complex in that here the two alternatives at once generate *error* as well as *doubt*, the former (error) representing a *failure in quick analysis* not strictly absence of knowledge. Suppose we analyze this situation in two pure doubts: (*yeh sarpa hai yā nahin hai*) ∨ (*yah rassī hai yā nahīn hai*). The complexity of the situation arises because the four alternatives are never *explicit* like in the earlier situations of pure doubt where only two alternatives obtain. The increased number of alternatives within a single perceptual experience on the one hand, and the failure of the mind to distinctly perceive them as alternatives on the other hand affect the genuineness of knowledge itself. Therefore, the exclusion awaits further inquiry which may or may not yield the result. That is to say, these alternatives too are *merely statements* of genuine knowledge not *propositions* of genuine knowledge thus requiring *siddhi* or falling short of *siddhi by utsargakaraṇa/anugaman*. If one were to proceed to examine the perceived situation in the manner of above analysis, one would first doubt whether, say, the alternative of presence of snake or of absence of snake is true; and then, when one of the alternatives is excluded, the others would be excluded automatically; but if not resolved, one would proceed to doubt whether the alternative of presence of rope or of absence of rope is true and then one of these would be excluded so as to settle the doubt finally. The fact, however, that the doubt is a *mixed* doubt disallows this pure-doubt-method of excluding alternatives *at that moment* due mainly to numerousness of alternatives. *Both error* and *doubt* therefore arise in such situations because human mind can *naturally* settle *instantly* only pure doubts thus not allowing the error to emerge but cannot *naturally* settle *instantly* the mixed doubts so that error emerges/props-up.

VII

The question whether *avaktavya* or unsayable is *apavāda* of sayable, or, exclusion or absence of sayable, is very important particularly for cosmological *saṁvāda*/inquiry. It is important to note that the *unsayable* generally refers to a specific experience not of perceived something but of *sākṣāt* or inner experience or seer-insight or ecstatic vision of cosmic something which although directly accessible to innermost seer, cannot be brought into linguistic usage/mode within the norms laid down by *Vyākaraṇa Śāstra*. Moreover, from logical point of view, the point of issue is not merely that of sayability but rather it is this that whatever is sayable must also be provable. Thus, the concept of *vaktavya* presumes that whatever is said is said logically, that is, proved. Not that all inner experiences cannot be said logically, for *some* of them can definitely be said and proved as in Sāṁkhya cosmology for instance. However, what seems to be the case is that after all that which could be said has been said, much more remains unsaid and that is where the significance of *avaktavya* lies.

Now, certainly by unsayable is not meant *silence*, that is, that which *could* be said but *was* not said. It rather means that which cannot in principle be said or depicted correctly and established,—that is, said in *vyākrata vāgbhāṣā* and established in *vyākrata saṁvāda*. It strictly means that some extraordinary inner experience has been had *non-linguistically* and therefore there is some awareness of 'something' which cannot be said at all logically. That is, not only can it not be defined or characterized essentially but also it cannot be *named* for if one could name it, one could say 'there is awareness of A (during *sākṣāt*)' and if one could characterize it one could say 'there is awareness that A is a (during *sākṣāt*)'. [A logically correct attempt to name it is *avyakta* and a logically correct attempt to characterize it is as *ahetumat* and *aliṅga* so that a logically correct attempt to say it would be '*Avyakta ahetumad hotā hai aur avyakta aliṅga hotā hai*'. However, upon analysis of this proposition-*rāśi* it will be found that hardly anything has been said.]

The inner experience seems to suggest that that *about* which one *wants* to say is excluded from that about which one *can* say. Therefore the *sayable* will be exactly this: *all* is said about that about which we *can* say. The unsayable then could not be that which *excludes* the sayable for the former is presumably not at all there, neither can it be the absence of sayable for this means absence of something which could be there but is not there, nor could it be total absence if we entertain the hope that it could be said some day in future, nor could it be *apavāda* of the sayable for the *apavāda* is also presumably there even if it is not said. Since the unsayable cannot be analysed in terms of absence or exclusion or *apavāda*, the only way left is to symbolize it as such. Now, we have symbolized propositions as ($p \cup q$), ($p \in q$) etc. but since there can presumably be neither terms nor

relations in the unsayable we can symbolize it only as (00). The question now is that how does it behave with respect to other propositions in different arguments? Can it be said to be true or false? Is it universal or particular or affirmative or negative? These questions cannot be answered satisfactorily at the moment but it is obvious that a thorough inquiry into logical features of (00) is called for. (00) seems to behave sometimes as true proposition and sometimes as false proposition. It never seems to *entail* any proposition but always *implies* some proposition and when so it seems to behave like a true proposition. Also when it occurs as exclusive disjuct, it seems to behave like a true proposition and excludes all other alternatives. Suppose a seer says '*yah vyādhi prāṇahara hotī hai aur* (00)' regarding some patient who presumably suffers from some serious ailment. From the above proposition-*rāśi* the only implication that can be had is that '*rogī mara jāyegā*'. Similarly if one asserts that '*kucha nahīn kahā jā saktā ki yah tārā hai yā . . . hai*'. What the assertion suggests is that amongst the disjuncts (00) $\vee t_x V \ldots$ the first one excludes the rest. If, on the other hand, we depict the sum and substance of the unsayable situation of inner experience as '*kucha avaśya hai parantu kahā nahīn jā sakatā*', then the complex statement/proposition can be analyzed as conjunct of an elementary statement '*yah koi na koi vastu hai*' and an *apavāda* statement of elementary type as '*Isa vastu ke bāre men kahā nahīn jā sakatā*'. The instance of unsayable situation can indeed be treated as exceptional since most situations are by and large sayable. The problem with such an analysis, however, is that the parent 'statement' can never be claimed to be a proposition since there is no *uddeśya* at all nor can it be claimed to be a universality of any variety since the 'experience' of '*koi na koi vastu*' is not universally accessible. A more satisfactory strategy regarding the unsayable could perhaps be one of identifying or defining 'all that is sayable' as 'All' or Universality as such and leaving the rest (the unsayable) as undefined or as *outside* the fold of universality itself. If we become clear about 'All' or 'universal' or *sarva* before hand, as above, then we limit ourselves only to the sayable in our *saṁvāda* but recognizing at the same time that there is something even more than 'All' about which nothing can be said. This seems to be the strategy adopted in Sāṁkhya Siddhānta.

Another way of analyzing the sentence '*Kucha hai jiske bāre men kahā nahīn jā sakatā*' could perhaps be to treat it as a proposition—a special proposition obtained by *śabda pramāṇa* calling for no reference—of which the *uddeśya* term is '*kucha*' and *vidheya* is '*jiske bāre men kahā nahīn jā saktā*' so that the *uddeśya* is *unidentified* and *vidheya* is *unsayable*. Such a proposition would then represent the other extreme of logical discourse or *saṁvāda*.

The questions of defining a logical-All and the distinction between *sat* and *asat* are central in *saṁvāda* so that, at least in a two-valued logic, we may not allow any simple propositions about more-than-All as well as about *asat*. That which is more-than-All cannot be named or defined or established so that any proposition regarding it is *impossible* for lack of

definite *uddeśya* as well as *vidheva*. That which is *asat* can be named but since its *utpāda* is not possible, nothing can be predicated of it, therefore, the only proposition regarding it would be '*Asat kā anutpāda hotā hai*', *asat* being a common name for *nṛśṛṅga, vaṇhyāputra, ākāśakusuma,* etc. Thus, if any property belonging to logical-All is bredicated of *asat*, the proposition would be meaningless. We therefore must ensure that both these types of propositions do not find entry in our *saṁvāda* and all our *uddeśyas* and *vidheyas* have definite meaning referring to *sat* only and neither to *ādhikya* nor to *asat*. It may here be asked if 'logical-All' itself is included in *sat*? The answer would be that the former is a more *definite translation* of the latter and the latter is a common name of *all* the objects of *samvāda* not itself being a subject of *samvāda* as a *rule*. Thus, if we utter a proposition such as 'logical-All as an entity belongs to logical-All', it would imply a fallacy of *anavasthādoṣa* within the system and would thus be inadmissible, meaning thereby that it has to be pushed into the realm of *ādhikya* or unsayable *within the system.*

Towards a Field Theory of Indian Philosophy: Suggestions for a New Way of Looking at Indian Philosophy

DAYA KRISHNA

Indian philosophy is usually treated in terms of the so-called six 'orthodox' and three 'non-orthodox' schools which are designated as Mīmāṁsā, Vedānta, Sāṁkhya, Yoga, Nyāya and Vaiśeṣika, on the one hand, and Buddhism, Jainism and Cārvāka, on the other. One may add a few more, but this is the usual way of presentation and it is taken as adequate by everybody. But, is it really so? Does it help us in understanding or grasping the philosophical scene in India as it unfolded over three millennia of its recorded existence?

Philosophers of diverse persuasions were not thinking in a vacuum or treating problems in monadic isolation from each other. The problem of understanding the Vedic text, and the Vedic ritual haunts all thinkers who had anything to do with them. The oldest Brāhmaṇa texts already evidence this concern, and so does the *Nirukta,* possibly the first text to deal with the problem. And so do the *Mīmāṁsā-Sūtras,* and the *Brahma-Sūtras,* though the latter deals primarily with the Upaniṣadic portion of the Vedic text. The problem, however, is a wider one and concerns the Buddhist and the Jain Canonical texts as well. The texts deemed 'sacred' have to be preserved, unaltered in every way and transmitted from generation to generation and carried to places far off from where the original was placed.

The purity in the preservation of a linguistic text, whether oral or written is, however, one thing and the understanding of what it means quite another. The two require very different sorts of skills and hence give rise to two different classes of specialists. Before the invention of tape-recording and writing, the task of preservation was so enormously complicated that it gave rise to the cultivation and development of specialized skills which are difficult to imagine today. We hear of persons who could remember after hearing just once and others who could do so after hearing only two or three or four times, and then we had persons continuously 'repeating' what they had heard so that what they had heard would not get lost.

That this activity of 'ceaseless repetition' was continued for millennia

is a miracle, but man as a species has been performing miracles continuously and this may be regarded as one of his defining features, even if it may not be the only one; once the 'miracle' is produced, it no more remains a 'miracle' and, if it is repeated often enough, it is hardly even thought or felt to be so. But the task of understanding what is so preserved requires a different kind of activity and a different kind of skill. But however different, it has to be as continuously undertaken as the former. This brings into being a specialized class whose whole activity consists in preserving the texts and interpreting them with a fairly clear-cut division between the two.

In India this class came to be known by the name 'Brāhmaṇa'. It consisted mainly of three different groups specializing in the task of preservation, interpretation and knowledge of the ritual according to which the Vedic *yajñas* were to be performed. The division is well known in the context of the Vedic tradition, but it must have obtained in the Buddhist and Jain traditions also, though one would have to find what the counterpart of the ritual of the *yajña* was in those traditions. Also one would have to find what the counterpart of the Vedic *Śākhās* was in those traditions, as the task of preservation and interpretation cannot be carried out for long without a specialized class of persons who are entrusted by the society for doing so and who are given sufficient honour, patronage and prestige to carry on the task they are entrusted with.

The task of understanding or interpreting a text has, however, an intrinsic 'undecidability' about it which has generally not been noticed until now. It was because of the essential ambiguity of language and the fact that it has not only literal but metaphysical meaning as well. There is also the problem of the 'uses' to which a language is put and the context in which it is being used. Similarly, there is also the presupposition of 'coherence' which imposes both an external and an 'internal' limitation on the interpretation.

The diversity of interpretation is thus in-built in the very task of interpreting as it never is in the case of the preservation of either the text or the ritual practice. One may, of course, make mistakes in these, but the 'mistakes' not only arise because of different factors, but are also of a different order. The variation in 'preservation' may occur because of not hearing properly what was said or the inability to decipher what was written because time or accident had obliterated something and different scribes had imagined what was to be inserted differently. This, as everyone knows, continues to be the situation in spite of all the advancements in the technology of preservation.

The problem in the preservation of the purity of exact instructions regarding ritual observances is of a different order as it emanates from the fact that no instruction manual, whether oral or written, can ever give all the details which are required for the performance of the action. These have to be learnt by actually observing the action performed and

by repeating it where every deviation is corrected by the master. This is one of the reasons for the importance that is accorded by the tradition to the 'Guru' as he alone can tell you the 'how' of the action and demonstrate it to you. The vocal preservation of the text involves this to a certain extent, particularly when it involves *svarānupūrvi* or even the singing of the *sāma*, as was the case with the Vedic texts. But the whole sequence and repertoire of a Vedic *yajña* was a totally different thing and its exact preservation became the prototype and the paradigm of the *prayoga śāstras* in the tradition.

The text known as *Brāhmaṇas* which are the primary manuals describing how to perform a particular *yajña* also discuss the possible meanings that the ritual may have. Dr. Mukund Lath has recently drawn our attention to this neglected aspect of the *Brāhmaṇa* text. Thus, the *vidhi* texts themselves have two different parts, relatively unrelated to each other; one relating to the performance of the ritual and the other to its interpretation. But these interests were not only different but divergent in character, as was shown by later developments. The *Mīmāṁsā Sūtras* carry on the task of filling in and accounting for the apparent discrepancies and inconsistencies in the instructions given regarding the actual performance of the various *yajñas*, while the *Āraṇyakas* and the Upaniṣads carry on the interpretative part not only independently, but even in opposition to them. The statement of Nārada in the *Chāndogya Upaniṣad* already indicates this, for he includes the four Vedas in the knowledge which he had acquired of the various disciplines known in his times though he remained unsatisfied and unfulfilled in spite of them.

The *Brahma Sūtras* thus ignore the *Brāhmaṇa* discussions almost completely and concentrate only on the Upaniṣads. The later commentators on the *Brahma Sūtras* and the Upaniṣads who wrote *bhāṣyas* on them by-pass the discussions of Yāska on the interpretation of the Vedic text in his *Nirukta*.

The *Nirukta*, in fact, has no direct successor and till almost the seventh century A.D., when we have the *Bhāṣya* on the *Ṛgveda* by Skandaswāmin and some others of the Vallabhi region of West India. But while there appears to have been no direct continuation of Vedic exegesis after Yāska for almost a thousand years, the concerns with the problems relating to the interpretations of linguistic meaning continued to exercise the Indian philosophical mind till today. *Śābda-Bodha* or *Vākyārtha-Vicāra* became one of the perennial concerns of all schools of thought, and one cannot understand the positions they take on the issue in isolation from those taken by others as if there were no relationship between them.

The same is true of the problem concerning the means and criteria of valid knowledge in the tradition against Nāgārjuna's questioning of the very possibility of there being any ground or criteria for such knowledge. It became incumbent for all to fight such radical scepticism. Even earlier, the Upaniṣads had raised the question as to how the knower itself can

even be known or as in the *Nāsadīya Sūkta,* how the question of ultimate beginnings could ever be answered. The author of the *Mīmāṁsā Sūtra* had argued that the knowledge of *dharma* or 'what ought to be done' cannot be derived either from perception or inference and hence required a third independent source for its knowledge. The *Brahma Sūtras* on their part explicitly declared that 'tarkar' has no place in the knowledge of Brāhmaṇa.

But all these have restricted the role of the *pramāṇas* to a certain realm, usually that of *vyavahāra* and not denied its relevance altogether. It is only Jayarāśi, the extreme sceptic, who is supposed to have done that. Matilal in his book entitled *Perception* appears to have treated his position as the standard one and had seen the whole of Indian philosophy as an attempt to meet the challenge. But there are degrees and variations of scepticism and to restrict the application of a particular one, or even to exclude *all pramāṇas* from application to a certain privileged one is not to deny the relevance of *pramaṇa* altogether as the debate between the votaries of *pramāṇa, samplava* and *pramāṇa vyavasthā* shows. It should also not be forgotten in this connection that *anubhava* or *sākṣātakāra* itself is a sort of *pramāṇa* and the authority of all *āpta vacana* is based on that. As in the case of *śābda bodha,* so also in the case of *pramāṇa vicara,* there is a host of problems cutting across various schools. The different schools are responding to a common concern, and unless we become aware of the common concern we cannot appreciate or understand what they are trying to do.

Take the notion of *duhkha* or 'suffering'. What exactly is the notion and why is there such a persistent and prevalent concern with *atyanta duhkhanivṛtti,* or the complete cessation of the very possibility of all sufferings for all times? This will involve a discussion of possibility and what it means. One solution is the state of the soul where it loses the very possibility of being conscious. But then what is the difference between such a soul and an inanimate object except that it is still characterized by having had such a consciousness in the past which the inanimate object can never be said to have had. Also, the difference of such a position with that of the Cārvāka would only be something like postulating a soul to survive after the destruction of the body. But, then, even for the grossest materialist something survives and it is generally called 'matter'. The only difference between this matter and the soul would be that the latter is still supposed to have a uniqueness of its own which the surviving matter lacks. But then what could this uniqueness consist of except with reference to the past which is now no more? This is something similar to what Strawson has discussed in his book entitled *Individuals.* But, then, such a reference to the past would also be there in the case of whatever matter is regarded as surviving after the destruction of the whole which constituted the body.

The problem of the whole and the part, of divisibility and indivisibility,

of the atom and the *ātman* are usually considered to be different and treated separately in the tradition. However, once it is seen that they are interrelated and the way one deals with any one of them has important consequences for the others, one would begin to see Indian philosophy in a different way.

There are so many other facets of Indian philosophy which would appear in a very different light if seen in this way. Take, for example, the problem of the ultimate constituents of the world as given in human experience. The Sāṁkhya, the Vaiśeṣika and the Jain scholars have treated this problem in a prominent manner. But in the usual presentation they are treated as if they were not dealing with the same problem and had nothing to do with each other. Yet, the moment we see them as attempting to deal with the same problem, we begin to notice the commonalities and the differences. Also, once we see them as tentative and provisional answers, we do not feel bound to the specificities of their positions and can benefit from the insights of all of them. But this can happen only if we see Indian philosophy as an 'ongoing enterprise' and not as something which only has a past and no future as if all the potentialities have already been exhausted.

But even if one does not believe that the creative potential of Indian thought in the realm of philosophy has not been exhausted and its immense possibilities of development in different directions are still unexplored, one would first have to realize the nature of the philosophical problems and questions with which the Indian thinkers struggled for at least two and a half millennia of recorded history. Unless this preliminary work is attempted, we cannot even hope to understand what they were attempting to do, let alone carry on their unfinished enterprise in any meaningful way. We have made a preliminary attempt in this direction, but it is confined only to the problematic and conceptual structure of their thought about man, society, polity and law. But the logical, epistemological, ontological and axiological issues which they struggled with have still to be articulated. Not only this, we have to see them in their development aspect, the way in which they developed over time.

To give but one example, we have to ask ourselves what was exactly meant by the terms Veda and *vidyā, śāstra* and *darśana,* and how these terms were related to each other. The term Veda has now acquired a sacred meaning, restricted only to the four *Saṁhitās,* but in ancient times it had no associations, as is evident in the use of such terms as '*Āyurveda*', *Dhanurveda, Nātyaveda* etc. The *Chāndogya Upaniṣad* mentions the four Vedas along with the other '*vidyās*' that Nārada had learnt and which had left him unfulfilled. The knowledge of the *Saṁhitās* was not given any special shape and was not included in what the Upaniṣads called the '*parā vidyā*'. The Mīmāṁsā tradition explicitly included the Brahmaṇa texts in the term Veda and treated them as equally authoritative.

The inclusion of *vārtā, daṇḍanīti* and *anvikṣiki* amongst *vidyās* raised

the question as to what the human end was that these *vidyās* served. Dr. Mukund Lath has recently drawn our attention to an interesting discussion by Udyotkara on this point. In his *vārtika* on the *Nyāya Sūtra,* 1. 1. 1. he says that every *vidyā* has its own *prayojanā* and '*niḥśreyas*' and specifically mentions *vārtā* and *daṇḍanīti* in this connection. According to him, only *ātmavidyā* has *mokṣa* as its *niḥśreyas.* There is an interesting discussion regarding this move of Udyotkara by Jayanta and Vācaspati Miśra I. But there appears to have been hardly any attempt to articulate and critically evaluate the different moves made on this issue in the tradition. Dr. V. Shekhawat has recently made an attempt to do this and has seen the whole of India's cognitive enterprise in a unique, developmental manner on the basis of the twin notions of 'models' and 'paradigms' derived from the recent work on the history of science by Kuhn and others.

The problem is the same with many other issues in Indian philosophy such as, say, the one relating to '*pramāṇa samplava*' or '*pramāṇa vyavasthā*', or, '*svatah pramāṇatva*' vs. '*paratah prakaśatva*' or, '*abhitānvayavāda*' vs. '*anvitābhidhānavāda*'. There is, in fact, no problem the discussion of which does not end across the boundaries of a traditionally demarcated system. Yet, even though the fact is well known, its radical implications for the understanding, comprehension and presentation of philosophical thinking in India has not been seen. It is not the problems and the issues that are seen as central and the *siddhāntas* of the so-called schools as peripheral, but instead the latter are viewed as central and the forms as peripheral. This is because the self-identification of the thinker is treated as more important than the problem he is concerned with. But, philosophically viewed, it is the latter that is important and not the former. The question, for example, of what is meant by '*śruti*' and what is to be regarded as such is more important then the specific answer that a Mīmāṁsaka or a Vedāntic or a Naiyāyikā or even a Buddhist or a Jain thinker gives to it. Once the situation is seen in this way, the distinction between the so-called '*āgama*' and the '*śruti*' will be seen as relating to the specificities of what is to be regarded as the foundational, authoritative text for a tradition. Similarly, the discussion about '*sarvajñatā*' and '*apauruṣeyatva*' would be seen in a new perspective and the question as to how the author of the *Nyāya Sūtras,* who seems to reject both, can still maintain the notion in his system. The example of *Āyurveda* given by him cannot be accommodated. Under neither case, does the bringing of Îśvara solve the problem. At least the author of the *Nyāya Sūtras* does not appear to treat him as the author of the Vedas which seem to be regarded as '*śruti*' by him. In fact the text *rejects* both the *Mīmāṁsā* and the *Brahma Sūtra's* interpretation of the so-called *śruti* as it nowhere talks of the *yajñas* and criticizes the view that all reality is one which it probably ascribes to the Upaniṣads.

It is, therefore, imperative that we get out of the prison-house of systems and focus attention on the problems, issues and questions that

troubled philosophers in India through the ages and the way they grappled with them and the arguments they gave for tentative answers and solution to them. Only through some such effort will we be able to enter into their philosophical world and see the inner, motivative force of the philosophical enterprise they were engaged in. And once this happens and we internalize their philosophical enterprise, making it an integral part of our own intellectual life, we will no more talk *about* it, but *do* it ourselves. Doing philosophy as an ongoing enterprise of philosophizing is *different* from talking *about* philosophy and we still have to learn the lesson, both about philosophy in the West and India, for unless we do that, we will always be 'outsiders' to both the traditions, making hardly any contribution to either.

Recent Theories of Truth

ANDRÉ CHAPUIS
Department of Philosophy, Indiana University, Bloomington
Indiana 47405–2601, USA

I. PARADOXES GENERATED BY THE NOTION OF TRUTH

Aristotle expressed our intuitive understanding of truth in his *Metaphysics* as follows:

> To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true.[1]

Although this formulation is very clear and seems unproblematic it gives rise to a problem that was discovered during Aristotle's lifetime: the Liar Paradox. This paradox is sometimes illustrated by a passage in St. Paul's *Epistle to Titus* where he attributes the following to Epimenides,[2] a Greek sage living at the beginning of the sixth century BC:

> One of themselves, even a prophet of their own, said, The Cretans are always liars, evil beasts, slow bellies. This witness is true. [King James' version]

A paradox arises from this statement if we assume that, with the possible exception of Epimenides himself, all other Cretans indeed utter falsehoods only. The assumption that Epimenides' claim is true leads to the conclusion that he must be a liar, that is, what he says is false. If on the other hand he utters a falsehood, just like all other Cretans, then indeed all Cretan utterances are false, so what he says is true.

Interesting difficulties arise also from the sentence that claims its own truth: the so-called Truth-teller. If we assume it to be true then what it says is correct, so it is true. However, if we assume it to be false it is incorrect to say of itself that it is true, so it is false. In this case we don't arrive at a contradiction, but we seem to have equally good reasons to believe the sentence to be true or false. Since in both cases self-reference is involved, a tempting move might be to dismiss these examples as not meaningful. Yet, there are grave problems with such a 'solution'. First, there are many unproblematic uses of self-referentiality as in 'This sentence is written in English.' It is unproblematically true. Even Epimenides' paradoxical statement may become unproblematic. If we drop the assumption that all other Cretans are liars, that is, if one Cretan tells the truth, then

Epimenides' statement becomes simply and unproblematically false. There are other examples of circular sentences involving the concept of truth that have unproblematic solutions and are clearly meaningful.[3] We therefore cannot dismiss circularity so easily. We want to distinguish problematic from unproblematic uses of circularity. We want to understand why the problems occur. Ideally we would want a systematic and unified account of the unproblematic as well as the problematic cases. It should illuminate the puzzling cases by explaining convincingly why these sentences behave the way they do. Theories that aren't simply *ad hoc* to explain the peculiarities of truth, but that are general enough to cover both, the regularities and the peculiarities, have flourished during the past thirty years. In what follows I will survey several of these new and intriguing approaches trying to sort out their presuppositions, their similarities as well as their differences.

Among the authors discussed I distinguish three types of theories: context-sensitive theories, fixed-point theories, and revision theories. Sections III–V present an overview of these types of approaches. To se the stage for these contemporary theories it will be useful to review some key elements of the most important contribution to the theory of truth in this century on which all current theorizing is based: Alfred Tarski's work on truth.

II. TARSKI'S THEORY OF TRUTH

Tarski's concern is to give a definition of the notion of truth which is *materially adequate* and *formally correct.* To be adequate means to capture the meaning of the familiar notion rather than defining some new notion. The familiar notion Tarski takes as the basis of his discussion is the Aristotelian conception. Applying this conception to the sentence 'snow is white' yields the famous example:

> The sentence 'snow is white' is true if, and only if snow is white.

The quotation marks are used in this example to form a name of the sentence 'snow is white', the reason being that grammatically a name is required in this place, and that ascribing a property to an object is done by using the name, not the object itself. Thus, generalizing the above example, using '*p*' to stand for a sentence and '*X*' for its name (formed by the device of quotation marks as in the example, or by some other device) we arrive at equivalences of the form:

> (T) *X* is true if, and only if *p*.

These are called 'equivalences of the form (T)', or sometimes 'T-biconditionals'. Tarski's criterion for a materially adequate definition of the concept of truth is that such a definition should entail all equivalences of the form (T). It is important to note that neither the expression (T)

which is just a schema and not a sentence, nor any particular sentence obtained by replacing the variables in the schema defines truth. But these replacements 'may be considered a partial definition of truth, [...]. The general definition has to be, in a certain sense, a logical conjunction of all these partial definitions.'[4]

Suppose we allow a language to formulate a sentence *l* saying of itself that it is not true. So *l* is a name for the sentence '$\neg Tr(l)$'. The T-biconditional for this sentence is:

> *l* is true if and only if $\neg Tr(l)$.

Thus we obtain $Tr(l) \equiv \neg Tr(l)$, that is, *l* is true if and only if it is not true which is obviously contradictory. Analyzing the assumptions that have led to this contradiction, Tarski makes the following points:

(I) We have implicitly assumed that the language in which the antinomy is constructed contains, in addition to its expressions, also the names of these expressions, as well as semantic terms such as the term *'true'* referring to sentences of this language; we have also assumed that all sentences which determine the adequate usage of this term can be asserted in the language. A language with these properties will be called *'semantically closed'*.

(II) We have assumed that in this language the ordinary laws of logic hold.

(III) We have assumed that we can formulate and assert in our language an empirical premise such as the the statement 2 [2 specifies that *'l'* names the Liar sentence $\neg Tr(l)$] which has occurred in our argument.[5]

Calling (III) 'inessential' since the paradox can be reconstructed without its help, Tarski proposes to reject (I) or (II). And since changing ordinary logic would have many undesirable consequences, he chooses to reject (I), that is, not to use any *semantically closed* language. This is the motivation to introduce two languages, one to which we want to apply the concept of truth, the *object-language*, and one in which we formulate the definition of truth and which we use to talk about the object-language, the *meta-language*. We may of course want to talk about the meta-language as well and about the sentences true in it. To do so is tantamount to making it an object-language, which means that we have to use a language that relative to it is a meta-language. In this way, if we wanted to talk about still further meta languages, we arive at a whole hierarchy of languages in which each is a meta-language of a higher level than the one it is taking as its object of investigation.

A concrete example shall illustrate the approach. Let L_0 be a very simple language containing only the names *a*, and *b*, and a predicate *F*. Let *F* apply to *a* only. We can now evaluate sentences of this language. *Fa* (read: *a* is *F*) is true, *Fb* (*b* is *F*) is false, $Fa \rightarrow Fb$ (if *a* is *F* then so is *b*) is false, and

so is $\forall x Fx$ (everything is F). Suppose we want a formal language to be able to express these verdicts. So we construct L_1 which contains a truth predicate Tr_1 and it has names, say quotation names, for the sentences of L_0. We can now write Tr_1 ('Fa') $\neg$ Tr_1('Fb'), etc. to express 'Fa' is true, 'Fb' is not true, etc. It is obvious that for example the sentence 'Tr_1 ('Fa')' itself is true, yet this has to be expressed not within L_1 but rather L_2 as Tr_2 ('Tr_1('Fa')'). If we did allow L_1 to express the truth of a sentence within the language of that sentence we could produce the Liar paradox as follows. Suppose we add the name l to our language L_1 denoting the sentence $\neg Tr_1(l)$. And suppose that l belongs to the extension of Tr_1, that is the objects Tr_1 applies to. Then $Tr_1(l)$ is true. Expressing this within the language would yield Tr_1('$Tr_1(l)$'). However, since l = '$\neg Tr_1(l)$' we would also assert Tr_1('$\neg Tr_1(l)$'). So we assert of a sentence and of its negation that it is true. If you suppose that l does not belong to the extension of Tr_1 then '$Tr_1(l)$' is false and hence '$\neg Tr_1(l)$' is true. If we express this again with our truth predicate we would assert $Tr_1(\neg Tr_1(l)')$. In parentheses we have again a name for the Liar sentence, so by substituting l for it we would be asserting $Tr_1(l)$. These contradictions are avoided following Tarski's proposal. I will illustrate this for the assumption that l belongs to the extension of Tr_1. The case for l not in the extension of Tr_1 is resolved analogously. The truth of the sentence $Tr_1(l)$ has to be expressed in L_2 rather than L_1 by the sentence Tr_2('$Tr_1(l)$'). Using the quotation name for l in $Tr_1(l)$ yields again Tr_1('$\neg Tr_1(l)$'). But now the contradiction disappears. We are no longer claiming of a sentence and its negation that it is true. We claim that a sentence is $true_1$ and its negation is $true_2$.

It is worth pointing out that semantically closed languages, as defined by Tarski, do not necessarily lead to inconsistency. This has been discovered by Anil Gupta who proved in his essay 'Truth and Paradox'[6] that we may add a truth predicate for a language to it, provided that the syntactic resources are weak. Yet, they may still be strong enough to contain the means Tarski formulated in (I). A precise definition of exactly what resources may be allowed without inviting paradox is still an open problem. However, for languages sufficiently rich in resources the problem Tarski observes applies and the problem before us is to give a definition of truth for a language that doesn't entangle us in paradoxes.

Without going into the technical details of Tarski's definition of the expression 'true sentence' here[7], I will just sum up the result: Tarski gives truth conditions for all sentences of an object-language. His definition does entail all T-biconditionals for that language. Yet, the price to pay is that we may not allow the object language to have its own truth predicate. The truth predicate is part of the meta-language. Similarly for the truth predicate of the meta-languages, it is part of a meta-meta-language, etc. The resulting hierarchy of languages, sometimes called a 'Tarskian hierarchy', has met opposition in recent writings as a description of the working of truth in natural language. To be fair to

Tarski it should be said that natural languages were not his concern. He states explicitly that he doubts a definition of truth for our language of everyday life is possible, believing that the concept of truth applied to colloquial language leads 'inevitably to confusions and contradictions'.[8] What he offers is an analysis of the problem and a solution for *formalized languages* which may be seen in analogy to the type theory of Russell. More recent authors have placed more emphasis on the natural language and its apparent capacity to contain a truth predicate for itself. Thus they are dissatisfied with a Tarskian hierarchy. The alternatives they offer and to what extent they succeed is the subject of the following sections.

III. CONTEXT-SENSITIVE THEORIES OF TRUTH

The distinctive feature of context-sensitive theories of truth is a parameter in the context of the utterance which explains, so it is claimed, the peculiar behaviour of truth. In the theories surveyed here this parameter is the context-sensitive extension in Burge's approach. For Barwise and Etchemendy it is the 'situation' an utterance is about. Gaifman's theory does not invoke an 'extra-linguistic' context, yet he shares in particular with Burge a preoccupation with what I call the 'Chrysippus intutition'. This is the idea that in the utterance of a Liar sentence something has 'gone wrong' which can be truly expressed by the same type of sentence. It becomes then necessary to explain why in certain contexts the sentence goes wrong and in others the seemingly same sentence doesn't. Thus Gaifman is included in this section. To illustarate this idea I go through the example from which the Chrysippus intuition derives its name. Aristotle says: 'What Plato says is true.' Plato says: 'What Aristotle says is not true.' If this is all they say it is a Liar cycle, that is, it works like the Liar paradox. Chrysippus overhears them, notices the paradoxicality, and he says: What Aristotle says is not true. The Chrysippus intuition is that what Aristotle and Plato said is paradoxical, but what Chrysippus said is true, even thought it appears to be the same as what Plato said.

Burge's indexical view

Tyler Burge aruges that the truth predicate changes its extension depending on the context it appears in. The example he uses to motivate this view is a variant of th Chrysippus example:

> Suppose a student, thinking that he is in room 10 and the teacher in room 9 is a fraud, writes on the board at noon 8/13/76: (*a*) 'There is no sentence written on the board in room 9 at noon 8/13/76 which is true as standardly construed'. Unfortunately, it being Friday the 13th, the student himself is in room 9, and the sentence he writes is the only one on the board there-then. The usual reasoning shows that it cannot have truth conditions. From this, we

> conclude that it is not true. But thisleads to the observation that (*b*) there *is* no sentence written on the board in room 9 at noon 8/13/76 which is true as standardly construed. But then we have just asserted the sentence in question. So we reason (*c*) that it is true.[9]

In the moves from (*a*) to (*b*) to (*c*) the linguistic meanings of the expressions involved don't change. But since the evaluation of the sentence written on the board changes, Burge concludes that some indexical element must be at work. 'True' is the only plausible candidate and thus he writes: 'The central idea in accounting for the move from (*b*) to (*c*) will be to interpret 'true' as contextually shifting its extension'..[10]

Burge's claim is not that there is a multitude of different truth predicates, as in a Tarskian hierarchy. Truth is a 'schematic predicate', he says, that takes on a specific extension in a given context. In that context we can represent 'true' with Tr subscripted numerically. Compare it to the indexical 'here' or 'now'. Each has a 'schematic meaning,' we know what it means in general. But the actual meaning, the place or time it refers to, is determined in a particular context of use only. Likewise for truth. Yet, the situation is more complicated. Statements involving the truth predicate may refer to other statements involving the truth predicate. This may continue and create chains of statements which may finally lead to a 'truth-free statement'. It may also create circles, however, or never ending chains. The numerical subscripts are a way of keeping track of the dependencies of different occurrences of 'true' in a given context.

Burge also uses the subscripts as a way to distinguish the unproblematic from the pathological uses. The truth-free, or so-called non-semantical statements are true_1 or false_1. Building on these we get higher and higher levels just as in the example illustrating the Tarski hierarchy. All semantical statements that are built on a prior level are called *rooted.* These are considered to be the unproblematic. Pathologicality, in contrast, consists in rootlessness. Having a root does not necessarily mean that a semantical statement is ultimately grounded in a non-semantical statement. These cases are indeed unproblematic. Burge explicitly expands the unproblematic 'to account for intutitions about the move from (*b*) to (*c*)'[11] in the example above. That is, a statement about a pathological statement is rooted, namely in the pathological statement itself. Furthermore he lays down an axiom[12] which implies that 'rootless sentences are not true.'

This is what makes Chrysippus' judgement come out true, as desired by the author.

Some problems

First a methodological point. It seems unsatisfactory to start with a notion of pathologicality to construct a theory that is supposed to explain this very phenomenon, that is pathologicality. It is even more unsatisfactory if the guiding intuition is not generally shared and fairly uncontroversial. Burge's rootedness as a mark of the unproblematic is controversial. We might for example argue that a pathological sentence infects all statements about it, making them equally pathological.

It would thus be more illuminating to hear how truth works in general, especially in the truly unproblematic cases, and then see from this account how the pathologicalities fit in Barwise's and Etchemendy's theory, as well as the revision theories, are better explanations of the phenomenon because they do precisely that.

A second point is related to the classification of phenomena that a theory allows, that is, the similarities and differences it allows us to see. Even though the Liar and the Truth-teller are both pathological they strike us as very different in nature. These different natures are not accounted for on Burge's account. A Truth-teller on his account is not rooted and thus pathological_i, that is pathological on some level *i*. That makes it not only untrue_i, but untrue_k for $k > i$. This is the natural verdict for one of the two possible assumptions. If we assume the Truth-teller to be false, then it is indeed false because it incorrectly claims to be true. The fact that the Truth-teller turns out to be true if we start with the assumption that it is true is lost however. Worse than that: Burge says that there is nothing wrong with asserting pathological_i sentences as long as the sentence is true_k, for some $k > i$.[13] By this standard it is okay to assert the Liar, but not the Truth-teller!

Consider, finally, the following sequence of sentences:

$1 \neg Tr(2)$
$2 \neg Tr(3)$
$3 \neg Tr(4)$
. .
. .
. .

Each claims its successor is false. If we assume that they are all subscripted alike, say with *i*, then all are pathological_i. Thus they are false_i. But expressing just this makes them true_{i+1}. What about being true_{i+2}? (1) is true_{i+2} about (2) saying $\neg Tr_i(3)$. But it is not true_{i+2} about (2) say in $\neg Tr_{i+1}(3)$. Now imagine all possible combinations of levels! We seem to be left with an enormous variety of possible assignments of different subscripts to all the sentences. This is where the pragmatic principles come in that should determine how subcripts are to be assigned.

Pragmatic principles

The principles in establishing the subscripts Burge laid down are explicitly formulated in the essay 'The Liar Paradox: Tangles and Chains'.[14] In the order of their importance[15] they are:

> *Justice.* Subscripts should not be assigned so as to count any given sentence substitutable in a truth schema instead of another, without some reason.

Verity: Subscripts on occurrences of 'true' (or 'satisfies') are assigned so as to maximize the applicability of truth schemas to sentences and minimize attributions of rootlessness.
Beauty: (Also *Minimalization*) The subscript on occurrences of the predicate 'true' (or 'satisfies') is the lowest subscript compatible with the other pragmatic principles.

Applying these principles to the list in the previous subsection yields that by *Justice* they are all assigned the same subscript.By *Verity* we want to maximize the applicability of truth schemas and minimize rootlessness. Minimizing rootlessness is impossible. Hence we assign by *Minimalization* the lowest subscript: 1. All sentences in the list are rootless$_1$, hence untrue$_1$, but they are rooted$_2$ and true$_2$. By a principle that Burge lays down truth is cumulative. That is, if a sentence is true$_i$ then it is true$_k$, for $k > i$. That means that the problem suspected in the list is greatly simplified.$\neg Tr_3$ (2) (as all the others) is thus true.But now that the sentence $\neg Tr(3)$ was just evaluated true 'on reflection', should we not evaluate $\neg Tr(2)$ false on further reflection? This is left unclear in both essays by Tyler Burge.[16] It seems that if we introduce simplifications,through principles like the proposed ones, we miss intuitively correct judgements. To get the judgements right the pragmatic principles would have to be complicated tremendously.

The Burgean theory of truth is thus somewhat deceiving in its simplicity. It is not simple because of an elegant solution to the difficulties, but because it relegates these difficulties to a pragmatic theory which is given only in a rudimentary way.

Gaifman's 'Pointer Semantics

In 'Pointers to Truth'[17] Haim Gaifman defends a theory of truth he has developed over a number of years. The technical core of his theory was published in 1988 in a paper entitled 'Operational Pointer Semantics: Solution to Self-Referential Puzzles I'[18]. In 'Pointers to Truth'[19] Gaifman outlines his technical apparatus giving us an intuitive understanding of how it works. But much of the paper consists of a clarification and defense of the philosophical underpinnings of the theory. Reversing Gaifman's order of exposition I will first address his philosophical premisses and then turn to the technical apparatus.

A major motivation for Gaifman's theory is to describe a mechanism he perceives in natural language which allows us to solve the following two-line puzzle:

line 1: The sentence on line 1 is not true.
line 2: The sentence on line 1 is not true.

The first sentence involves us in an unending cycle and it is thus, according to Gaifman, not true. The second sentence states this conclusion

and it is thus true. Since we have the same sentence (type) on both lines the puzzle is that this one sentence seems to be both true and not true. Gaifman resolves the case by assigning truth and falsity not to sentence types but to their tokens, allowing for the assessment of different tokens of the same type to differ in a limited way. His system is based on the three truth values T, F, and GAP ('neither true nor false'). Two tokens of the same sentenceare never T and F, but all other combinations are possible. For example, the token on line 1 in his example is assigned GAP, the one on line 2 is assigned T. The mechanism Gaifman perceives is that even though certain tokens go bad, or fail, in unfavourable circumstances, as the one on line 1, we can use a different token that recognizes this failure and expresses it successfully. For generality Gaifman prefers to use the term 'pointer' instead of 'token'. A pointer is any object used to point ot some object. A token is a special kind of pointer, it points to the type of which it is a token. On the pre-theoretic level there can be honest disagreement as to whether the sentence on line 2 is true. We might argue that this sentence fares no better than the one on line 1 when it comes to determining whether it is true or not. It might be said that the first sentence leads us immediately into a cycle whereas the second leads us first back to line 1 and then into a cycle. Thus both sentences resist straightforward evaluation.

The alternative view I just hinted at leads to a phenomenon Gaifman calls 'black holes', which he tries to eliminate as much as possible. Kripke, he claims, noted the problem but he didn't give the issue the central place it deserves. On Kripke's system not only simple paradoxes like the Liar (L) 'fail' but any attempt to predicate truth or falsity (TrorFa) of the Liar embedded in any sequence of predications of truth or falsity fail, for example, $Tr(L)$, $Tr(\neg(L)))$, $Fa(Tr(Tr(L)))$, etc. are all neither true nor false. Since any semantic judgement about any embedded L is prevented by L's gappiness Gaifman calls La 'black hole'. He claims that '[l]anguage usage involves a mechanism for solving the black hole dilemma. And if our goal is to model natural language, or a basic aspect of it, then the elimination of black holes—at least those of the simpler varieties—should be an overriding concern'.[20]

This 'methodological premiss' is plausible if we share Gaifman's intuition. Yet, if we don't we may see 'black holes' in language as just as natural as the ones in space which nobody wants to eliminate. The difference here lies in the fundamental attitude one has towards the paradoxes, whether they are viewed as an illness to be cured or a phenomenon to be understood. A second 'methodological premiss' is noteworthy: 'A semantics of formal languages—containing their own truth predicate—should do for the concept of truth what set theory, such as ZF, does for the concept of set. By minimizing the black-hole phenomena we are shifting the problematic points, w here the system's inadequacies show to the less accessible parts of the system'.[21] Pointer

semantics does not get rid of all black holes, they reappear if we increase the expressiveness of the language. But Gaifman hints at more sophisticated versions of his theory which would eliminate even more black holes, so his present proposal is to be understood as a first step only, intended to 'tidy up the more elementary levels of language'.[22] A third motivation for the elimination of black holes, in the context of mathematical languages, is the goal of a 'universal language—a self-contained system, one that includes its own semantics'.[23] On the surface it would seem that 'shifting problems to the less accessible parts of the system' conflicts with the goal of a 'universal language'. But may be this could be explained if the rather vague notion of a 'universal language' were made more precise.

Overview of the technical apparatus

Sentences are represented by networks of pointers. A pointer p to a complex sentence $A*B$ (* is an arbitrary sentential connective) gives rise to (calls directly) derived pointers $p1$ and $p2$ pointing to A and B respectively. If p points to $\neg A$ then $p1$ points to A and $p2 = p1$. A pointer to $Tr(q)$ or $Fa(q)$ calls directly q. Representing the pointers with nodes and the calling relation with arrows, we get a network of pointers which may feature loops of all kinds. There are three kinds of rules that are invoked by the algorithm Gaifman sets up to evaluate the pointers of such a network The *Standard Rules* handle basic sentences, that is, senteneces not involving the truth predicate—pointers to those are assigned whatever the model assigns—and complex sentences as long as the strong Kleene truth tables[24] allow a decision. Example: If p points to a conjuction thus calling $p1$ and $p2$ we can determine p's value if both conjuncts are either T or F or if at least one is F. Similarly for the other connectives. We make $Tr(q)$ true if q is, false if q is false. The *Gap Rules* determine the assignment of GAP. There are three of them: a closed loop, that is pointers calling each other such that there is a path from any one to any other, gets assigned GAP (Closed Loop Rule). A pointer whose derived pointers have been assigned a value but which cannot be evaluated by the Standard Rules is assigned GAP (Simple Gap Rule). In the example above this would be invoked if one conjunct is T and the other GAP. Finally, if all other rules cannot be invoked and there are still pointers left to evaluate, the remaining pointers are assigned GAP (Give-up Rule). The third kind of rule, the distinctive feature of pointer semantics, is the *Jump Rule*. If p points to $Tr(q)$ or $Fa(q)$ and q but not p is assigned GAP then the Jump Rule assigns F to p. Intuitively, if q is GAP, it is false to say that it is true (or false). Hence p, which does just that, is false. The Jump Rule allows the judgement Gaifman wants to make in the two-line puzzle. Line 1 is evaluated GAP because it is a closed loop; line 2 gets T, the Jump Rule assigning F to the unnegated sentence and the Standard Rule for negation changing that to T. In general we start with a network of unevaluated pointers and then apply the rules which can be invoked to assign more and more truth values to the pointers. The procedure guarantees that we obtain a complete asssignment of semantic values.

It takes working through a few examples to get a better feel for the system, which is not possible in this limited space. But the brief outline is, I hope, sufficient to make the following general remarks intelligible. Pointer semantics is a departure from traditional compositional semantics in the following sense: 'The meaning of a [sentence] token is not reducible to what the sentence says, even when the denotations of its terms have been fixed. It derives also from the token's particular place in a global set up'.[25] This complicates semantics and is part of the price we pay. What we get in return is the following. In Gaifman's words, 'Tarskian semantics or any of its offshoots', including Kripke, Gupta, Herzberger, cannot accommodate an ascent in the linguistic hierarchy, that is, we are forced, by the paradoxes, to expand our language and use a higher level predicate 'true,' or a new predicate like 'stable truth'. Pointer semantics, in contrast, encompasses such an ascent in a single language. Yet, whether the peculiar phenomena caused by the truth-predicate that leads some to introduce new predicates ought to be described by the truth-predicate itself or a different predicate is a philosophical question that is not easily answered, and it is not addressed by Gaifman. It is unclear, for instance, whether English is 'a single' language in the sense intended in pointer semantics. For more on these issues see section VI.

Summary

Gaifman perceives a mechanism in natural language that allows us to recognize failures of evaluation as they happen in pathological cases. He sets up an evaluation algorithm that reflects the recognition of failures, the noticing of the pathologicality, by declaring sentences true (false) that correctly (incorrectly) notice the pathologicality of the sentence(s) they are about. This is a new way of looking at evaluations of sentences and it can be adapted to work with other presuppositions than Gaifman has, as for example the three-valued logic he employs, or viewing sentence tokens as bearers of truth. The context he invokes is that of a network of sentence tokens in which a given sentence is to be evaluated by this algorithm.

Gaifman and Burge both think that pathologicality should be expressible with the truth predicate. This idea by itself is controversial. Yet, it explains to a large measure the shape their accounts are taking. They differ sharply with respect to the role pragmatics is supposed to play. It is the key to understanding the peculiarities of the concept of truth according to Burge. On Gaifman's view, in contrast the key is the 'network creating nature' of the predicate 'true' which explains the difficulties.

Barwise's and Etchemendy's Theory

The theory of Barwise/Etchemendy is in many ways more satisfying than the ones presented so far. It is not fixated on one single intuition, that is, the Chrysippus intuition. It is based on a general conception of truth gathered from John Austin, and an account of propositions embodying this notion of truth. As a result from these general presuppositions about truth and propositions we get an account of what is pathological.

Two aspects of their conception of propositions are important. One is the circularity they allow, that is, propositions may be about themselves. To make sense of such propositions in a mathematically precise way they model their propositions with the non-well-founded sets of Peter Aczel's set-theory.[26]

The second is a situational parameter entering the proposition.The second aspect is the more important for them, it reflects their philosophical view of truth. Their view is that propositions should be relativized to situations. Situations make propositions true by containing the appropriate facts.Or: a proposition is true if the situation it is about is of the type the proposition claims it is. Example: 'Snow is white' is true if the situation this proposition is about is of the type that snow is white. The actual situation we find ourselves in when we use that sentence is of that type (by and large), and so the proposition is true. Thus they give an account of what they call *Austinian propositions* that feature situations as components.

Austinian truth and propositions

Austin's notion of truth is best summarized in his own words:

> A statement is said to be true when the historic state of affairs to which it is correlated by the demonstrative conventions (the one to which it 'refers') is of a type with which the sentence used in making it is correlated by the descriptive conventions.[27]

Using the term 'proposition' instead of Austin's 'statement,' Barwise/Etchemendy call a proposition p true if the situation it is about (s) determined by the demonstrative conventions is of the type (T) determined by the descriptive conventions. In short: A proposition p is represented by a set-theoretic object $\{s; T\}$, and it is true if s is of type T. The basic building blocks to describe the world are states of affairs. They are built of individuals having certain properties or standing in certain relations. A set of such states of affairs is a situation. A state of affairs also determines a type. A situation is thus going to be of all the types of the states of affairs it contains. Hence we may also say that p is true if and only if the state of affairs that determines T is part of s.

On the Austinian account the truth of a proposition seems so far to depend only on situations and what states of affairs they contain, and not on the world. The role the world plays is in the distinction between

'accessible' and 'inaccessible' propositions. If a proposition is about an actual situation it is called 'accessible', if it is about non-actual situation it is 'inaccessible'. This distinction is not related to the truth or falsity of a proposition. There are true and false propositions about actual and non-actual situations. Here are a few examples: let (snow, white) be the state of affairs of snow being white, and let us use square brackets to indicate the type. So $sw_1 = \{(\text{snow, white}), [\text{snow, white}]\}$ is the proposition 'snow is white' about a very restricted situation, one containing only the fact that snow is white. Since this situation is of the type [snow, white] it is true. And since the state of affairs (snow, white) is part of the world the proposition is accessible. Now consider the variation $sw_2 = \{(\text{grass, green}), [\text{snow, white}]\}$. The situation the proposition is about is not of the required type, and thus the proposition is false, yet still accessible. An inaccessible proposition would be one that is about, for example, the state of affairs (moon, cheese). If the proposition 'the moon is made of cheese' ($mc = \{(\text{moon, cheese}), [\text{moon, cheese}]\}$) is about a situation containing that state of affairs it would be true.But it would of course be false about a situation that did not contain this state of affairs. Given that (moon, cheese) is not part of the actual world any accessible proposition claiming that the moon is made of cheese would thus be about situations that are not of the required type and would thus be false.

This is also what the authors want to say about the Liar f_s if it is taken to be about some actual situation s. $f_s = \{s; [Tr, f_s; 0]\}$ is simply false about s.

Yet, there are true Liars, that is, those about situations that contain the fact $(Tr, f_s; 1)$. However, those situations they are about cannot be actual. The Truth-teller $t_s = \{s; [Tr, t_s; 1]\}$ is true about some actual situations false about others.

The intuitive Liar reasoning that leads us through the switches in truth values is captured on their picture by a sequence of propositions as follows. The Liar f_{s1} about an actual situation s_1 is false, but we can expand s_1 to s_2 which includes the fact of the falsity of f_{s1}.The proposition $p_{s1} = \{s_2; [Tr, f_{s1}; 0]\}$ is true because it says about s_2 that it is of the type of the Liar about $s1$ and sure enough s_2 is of that type since it contains $(Tr, f_{s1}; 0)$. p_{s1} is not a Liar for s_2, but we can construct one, fs_2, and it is already clear that by the same reasoning as before we get a sequence of propositions f_{s1}, ps_1, f_{s2}, p_{s2}, f_{s3} . . . alternating in truth value and mirroring the flip-flopping behaviour of the Liar.

We thus do not see the switching of *the* Liar properly, because each one is false and stays false as long as it is supposed to be about an actual situation. What we see is that from a situation in which the Liar is false we can get into a situation in which we recognize its falsity with a proposition that is of the same type. In *that situation*, however, that proposition is not a Liar proposition. The proper Liar proposition for that situation is again false, etc. What is switching, or, more accurately, changing, are the

situations we use to evaluate the Liar type propositions.

In other words, the fact of the falsity of Liar about *s* diagonalizes out of *s*. Yet, although the situation cannot contain that fact, the model can, as part of a different situation. This provides them with a way to mirror the intuitive Liar reasoning we just saw. The Chrysippus intuition can be accommodated the way Gaifman and Burge want. When Chrysippus utters the seemingly same proposition as Plato he has a different reference situation, thus his proposition can be true, whereas those of Plato and Aristotle are false.[28]

In sum, the approach by Barwise and Etchemendy attributes the difficulties to ambiguity, the ambiguity with respect to which situation a propositions is about.[29] Whether ambiguity is really at work in the Liar is questionable. There is an argument to be made that ambiguity is not the key because of a feature of truth sometimes called the 'supervenience' of truth. The idea is that the non-semantic facts fix the semantic facts. If that is so, then ambiguity is not the source of paradox.[30]

The question is how to weigh certain intuitions. The intuition that some ambiguity is involved in the Liar that needs to be made explicit is certainly a strong one. The many 'ambiguity theories' of truth are a testament to that. Whether supervenience should have less, equal, or more weight depends on whether a plausible theory can be constructed that shows how non-semantic facts fix the semantic facts even in the pathological cases. The revision theories discussed in section V provide such a construction.

Robert Koons has provided an account that combines the situation semantics approach with the supervenience intuition. He shows how a homomorphism from Burgean into Austinian semantics can be constructed. To achieve this he has to restrict himself to expressible Austinian propositions, propositions featuring only actual situations as parameters, because inexpressible propositions cannot be handled by the Burgean framework. Conversely, the Austinian situations for which he can provide a Burgean model have to be appropriately complete. The 'fundamental completeness' of the situations he restricts himself to embodies the supervenience intuition.[31] Without going into the technical details of his approach what Koons provides can be summarized as follows. He shows how some of the differences between Burge and Barwise/Etchemendy can be eliminated by modifying each theory appropriately. The main change in Burge is the addition of 'inner' and 'outer' negation, two different forms of negation corresponding to Barwise/Etchemendy's negation and denial. Barwise and Etchemendy, on the other hand have to give up some of the rich structure of their Austinian propositions, that is, non-actual and semantically incomplete situations. Since Koons believes that there is no need to have more finely discriminating propositions than Burge's, he prefers the simpler account of Burge, combining it with as slightly modified Gaifmanian evaluation procedure.

The resulting theory claims as its advantages that it needs only two levels for truth predicates, rather than levels indexed by all ordinals as on Burge's view. However, the problems Koons faces are analogous to Burge's. If we rely heavily on pragmatic principles then explaining the workings of truth amounts to working out a theory of how these principles are invoked. In the absence of such a theory it is difficult to assess, for example, whether Koons' abandonment of the richer structure afforded by Austinian propositions is warranted.

IV. FIXED-POINT THEORIES

Here is in Kripke's words what he is after:

> We wish to capture an intuition of somewhat the following kind. Suppose we are explaining the word 'true' to someone who does not yet understand it. We may say that we are entitled to assert (or deny) of any sentence that it is true precisely under the circumstances when we assert (or deny) the sentence itself. Our interlocutor then can understand what it means, say, to attribute truth to (6) ('snow is white') but he will still be puzzled about attributions of truth to sentences containing the word 'true' itself. [. . .]
>
> Nevertheless, with more thought the notion of truth as applied even to various sentences themselves containing the word 'true' can gradually become clear. Suppose we consider the sentence,
>
> (7) Some sentence printed in the *New York Daily News*, October 7, 1971, is true.
>
> (7) is a typical example of a sentence involving the concept of truth itself. So if (7) is unclear, so still is
>
> (8) (7) is true.
>
> However, our subject, if he is willing to assert 'snow is white', will according to the rules be willing to assert '(6) is true'. But suppose that among the assertions printed in the *New York Daily News*, October 7, 1971, is (6) itself. Since our subject is willing to assert '(6) is true', and also to assert '(6) is printed in the *New York Daily News*, October 7, 1971', he will deduce (7), by existential generalization. Once he is willing to assert (8). In this manner, the subject will eventually be able to attribute truth to more and more statements involving the notion of truth itself. There is no reason to suppose that *all* statements involving 'true' will become decided in this way, but most will.[32]

Some of the statements that will not be decided are for example the Liar, or the Truth-teller ('This sentence is true'). The name 'fixed-point theory' is motivated as follows. Kripke shows how we can construct a

three-valued language containing its own truth predicate. We start with an interpreted first-order language L that is rich enough to allow a description of its syntax. We extend the language by adding a monadic predicate $Tr(x)$. The predicate is interpreted by a pair (S_1, S_2), where S_1 is called the *extension* of *Tr*, the set of entities to which *Tr* applies, and S_2 is the *antiextension*, the set of entities to which *Tr* fails to apply. For entities outside $S_1 \cup S_2$, $Tr(x)$ is undefined. Let the language $L(S_1, S_2)$ be the expansion of L obtained by interpreting $Tr(x)$ by the pair (S_1, S_2). $L(S_1, S_2)$ is a three-valued language and its sentences can be evaluated using any of the three-valued schemes, for example, Strong Kleene. Now let (S'_1, S'_2) be the pair consisting of the set of true sentences (S'_1) of $L(S_1, S_2)$ and the set of false sentences and non-sentences (S'_2) of $L(S_1, S_2)$. The language $L(S'_1, S'_2)$ can be thought of, intuitively, as a metalanguage of L that contains a truth predicate for L. What we are after is a truth predicate for the language itself, and to obtain one we must have that $Tr(x)$ agrees with what the meta-language declares to be the true and false sentences of L. This means, we need a language $L(S_1, S_2)$ such that $S_1 = S'_1$ and $S_2 = S'_2$. To put it differently: the function taking us from (S_1, S_2) to (S'_1, S'_2) must be shown to have fixed points. Kripke proves that this function he calls ϕ has fixed points for certain three-valued languages (for example Strong Kleene). That is, there are instances such that $\phi((S_1, S_2)) = (S_1, S_2)$. A fixed point means that we have found a language L such that the predicate $Tr(x)$ of L agrees with what a meta-language of L declares the true/false sentences of L to be. In other words $Tr(x)$ is a truth predicate for L.

Kripke's picture is attractive. There is an intuitive appeal in the thought that we decide about what is true and false in stages, getting better and better candidates. And it is surprising to learn that this process doesn't go on indefinitely but yields a 'best stage,' a fixed point which would seem an appropriate candidate for the interpretation of the truth predicate. We should note, however, that we can reach not just one but many different fixed points, depending on the initial interpretation of $Tr(x)$. Starting from the empty set, for example, we will arrive at a least fixed point in which all problematic or so-called 'ungrounded' statements have no truth-value. All unproblematic or 'grounded' statements will have their expected value. The grounded statements have the same truth-value in all fixed points, and there are some ungrounded statements that have no truth-value in any fixed point. For the latter Kripke suggests that we call them 'paradoxical.' Other statement are problematic without being paradoxical in the sense just defined. They may be true in some and false in other fixed points (the Truth-teller, for example), or true in some but not all fixed points. The fixed points which assign a statement a truth-value which doesn't conflict with the value it gets in any other fixed point is called *intrinsic* by Kripke. It can be shown that there is a largest intrinsic fixed point and Kripke underlines its attractiveness as the interpretation of truth: 'The largest intrinsic fixed point is the unique 'largest' interpretation of $T(x)$ which is consistent with our intuitive idea of truth and makes no arbitrary choices in truth assignments. It is thus an object of special theoretical interest as a model'.[33]

Yet, the least fixed point is certainly also a natural candidate for the interpretation of the truth predicate, as Kripke notes himself, and there are even other fixed points which have been proposed as candidates. In order to criticize the fixed-points theory based on concrete examples we would need to clearly commit to some fixed point as the preferred candidate, something that Kripke didn't do. Even though Kripke's paper may thus not be a proposal of *a* definite theory of truth it was and is an important stimulus for later theories of truth, and it has certainly fulfilled the hopes Kripke expressed in the paper, namely: 'I do hope that the model given here has two virtues: first, that it provides an area rich in formal structure and mathematical properties; second, that to a reasonable extent these properties capture important intuitions'.[34]

The nice property of reaching a fixed point fails when we use classical logic; the presence of a Liar, for example, prevents us from ever reaching one. Thus we cannot develop a fixed point theory of truth based on classical logic. Nothing prevents us of course from abandoning classical logic,[35] yet it is unsatisfactory to let the theory of truth decide about the logic we adopt. It would be preferable if our theory of truth were neutral in this respect. Although this wasn't the motive behind Hans Herzberger's experiment of combining Kripke's ideas with classical logic, it would eventually lead to just this result with the revision theory of truth of Anil Gupta and Nuel Belnap.

The Skyrmsian twist

Brian Skyrms's theory of truth starts from the following two sentences:

(1) (1) is not true.
(2) '(1) is not true' is not true.

His intuition is that (1) is neither true nor false and that (2) is true. (1) and (2) refer to the same sentence; the only difference is the way the sentence is referred to. This might suggest that the difference in truth-value is due to the difference in the way the reference is achieved. In other words, although the terms (1) and '(1) is not true' are coreferential they cannot be substituted *salva veritate* in the context '_____ is true'. Thus Skyrms concluded in 'Return of the Liar'[36]: 'I believe, then, that weakening of substitutivity of identity is part of the price that must be paid for a philosophically adequate solution to the liar paradox. This amounts to giving up Frege's principle that the denotation of a sentence (i.e., its truth values) is a function of the denotation of its constituents.' What Skyrms proposes instead in his paper 'Intensional Aspects of

Semantical Self-Reference'[37] is a variant of Kripke's Fixed-Point Theory of Truth. Skyrms adds some new twists which are needed to deal with the two sentences (1) and (2) above.

Intensional Interpretations

The truth predicate in Skyrms' system is not interpreted in the same way ordinary predicates are interpreted. The extension and antiextension of the predicate are called the 'D-tension' and 'anti-D-tension,' (D_1, D_2), of truth. D_1 contains only pairs of sentences and their names, D_2 contains in addition non-sentences and their names. Skyrms allows D_1 and D_2 to overlap. So a sentence may be both true and false. This is familiar from four-valued logics.

To evaluate a sentence $Tr(a)$ we look at the pair consisting of the name a and the sentence it denotes. If this pair is an element of D_1 the sentence is true. If the pair is in D_2 the sentence is false. The sentence is assigned the value **n** (neither) iff it is evaluated neither true nor false, that is, if it is neither in D_1 nor in D_2. And it is assigned **b** (both) iff it is evaluated true and false. The evaluation of complex sentences is based on an extended Strong Kleene scheme as follows. The negation of a sentence is true iff the unnegated sentence is false, and it is false iff the unnegated sentence is true. Note that being true or false does not necessarily coincide with being assigned **t** or **f**. A sentence assigned **b** is also true, as well as false. The above clauses determine however all possible assignments. If a sentence is neither true nor false then so is its negation, and if it is both true and false, then its negation is also both true and false. A conjunction is true iff both conjuncts are true, and it is false iff one conjunct is false. Similarly for disjunctions: A disjunction is true iff one disjunct is true; it is false iff both disjuncts are false. Quantifiers are taken substitutionally, that is $\forall x Ax$ is true iff all instances At are true, false iff some instance At is false. Since a universal quantification can be thought of as a conjunction of all its instances the assignment of semantic values functions analogously.

In his construction of fixed points, Skyrms maps a pair of D-tension and anti-D-tension to such a pair with a function g, which is analogous to Kripke's ϕ, mentioned above. g maps (D_1, D_2) onto (D_1^+, D_2^+) such that;

> (A) If x is in D_1 then x is in D_1^+. If x is in D_2, then x is in D_2^+.
> (B) If the interpretation of the language based on (D_1, D_2) makes x true and y is any name of x, then (x,y) is in D_1^+. If the interpretation of the language based on (D_1, D_2) makes x false or x is a non-sentence, and y is any name of x, then (x,y) is in D_2^+.[38]

Clause (A) assures that the D-tension and anti-D-tension only increase or remain the same but never lose members. This assures that we always reach a fixed point, because D_1 and D_2 couldn't keep increasing indefinitely. Clause (B) says that what is true (false) by some name is true (false) by any name. The function g lets us improve on a given candidate for a pair (D_1, D_2), the interpretation for the truth predicate. If we start

with the (ϕ, ϕ), say, we will reach the minimal fixed point in which, intuitively, D_1 contains pairs of grounded sentences and their names; and similarly for D_2. Pairs of circular sentences and their names, or sets of pairs forming circles, are neither in the D-tension nor the anti-D-tension of truth. Nothing will be in both D_1 and D_2 in the minimal fixed point.

But compare what happens if we start out setting the Liar either true or false. On Kripke's system the Liar would move in and out of the extension of Tr. This behaviour is different in Skyrms' system due to clause (A) in the definition of g. If we put the Liar and its name in either D_1 or D_2 at a stage, then at the next stage the pair is in both, and it remains in both through the repeated applications of g that lead to a fixed point. In this fixed point the Liar will of course also be in both D_1 and D_2. Unlike Kripke, Skyrms allows that the D-tension and anti-D-tension overlap, this is why he needs the four-valued scheme. He defines paradoxicality in terms of overlap as follows. A valuation is paradoxical iff the intersection of D_1 and D_2 is not empty. If a fixed point has overlapping D_1 and D_2 the valuation leading to this fixed point (by iterating applications of the function g) is called a *paradoxical valuation*. A *paradoxical sentence* is a sentence leading to a paradoxical valuation for any valuation which assigns it **t, f,** or **b**.

In Kripke's treatment the sentence $\neg Tr(a)$, where a is a name for that very sentence, is neither true nor false, and the same holds for $\neg Tr(b)$, using a different name b for $\neg Tr(a)$. There is no way to make $\neg Tr(b)$ come out true in a Kripke construction. If we put $\neg Tr(a)$ in the antiextension then $\neg Tr(b)$ will be evaluated true, but so also will be $\neg Tr(a)$ and we will not get a fixed point.

On Skyrms'treatment, in contrast, we can put $\neg Tr(a)$ by a different name than a in D_2. All that follows from $(\neg Tr(a), b) \in D_2$ is that at the next stage $Tr(b)$ by any name is in D_2 and consequently $(\neg Tr(b), c)$, is in D_1. So we now have the situation that $\neg Tr(b)$ is true, yet $\neg Tr(a)$ is neither true nor false.

In sum, Skyrms' construction does essentially what Kripke's does. We reach fixed points that seem plausible candidates for the interpretation of the truth predicate, but Skyrms doesn't decide on one particular such point. There is one minor and one major difference between the two approaches. The minor one has to do with the definition of 'paradoxicality'. Kripke suggests as a definition of 'paradoxical sentence' that they are those sentences which have no truth value in any fixed point. Skyrms defines them via overlap of D_1 and D_2, as mentioned above.

The major difference comes in the evaluation which in Kripke's theory embodies the full Tarski biconditionals. That is, on his approach p is true *if and only if* $Tr(p)$ is true. Skyrms, in contrast, has only 'half' of the biconditionals. That is, from having the pair (x,n) in D_1 we conclude $Tr(n)$, but we cannot go in the other direction, we cannot go from $Tr(p)$ to p.

This 'feature' is what makes Skyrms say that 'our [Skyrms'] construction never looks back'.[39]

Some problems

Skyrms' system has quite a few similarities with Gaifman's—Gaifman has different pointers instead of names that can express that something has gone wrong with some pointer—and these similarities seem to extend to the problems. One problem Gaifman has is that in an infinite chain producing a 'black hole' he cannot comment on its non-truth as he does in his two-line puzzle. Such a comment would be sucked into the black hole. A similar problem occurs for Skyrms with the quantifiers. Consider the sentence named a which is $\forall x\,(x = a \rightarrow \neg\, Tr(x))$. Intuitively $\forall x(x = a \rightarrow \neg\, Tr(x)$ is a Liar, it says of itself that it is not true. Placing $(\forall x(x = a \rightarrow \neg\, Tr(x)), a)$ into D_1 or D_2 will result in a paradoxical fixed point, as we would expect. However, the same happens if we place $(\forall x(x = a \rightarrow \neg Tr(x)), b)$ say in D_1. Doing this we would have $Tr(b)$ together with any name after one application of g in D_1 and hence we have $b = a \wedge Tr(b)$ and thus $\exists x(x = a \wedge Tr(x)$ with any name in D_1. Since $\exists x(x = a \wedge Tr(x))$ is equivalent to $\neg \forall x(x = a \rightarrow \neg\, Tr(x)), \forall x(x = \rightarrow \neg\, Tr(x))$ by any name is also in D_2 and we have a paradoxical valuation. The same can of course be repeated for any name.

Skyrms' solution has thus two shortcomings. On the one hand it cannot rule some sentences paradoxical that intuitively seem paradoxical. On the other his manoeuvre to recognize the paradoxicality of a sentence with a different name does not work in all cases, as is shown by the 'universal Liar.'

V. REVISION THEORIES OF TRUTH

Some background

The revision theories of truth have different philosophical motivations. Hans Herzberger (cf. his 'Notes on Naive Semantics'[40] was after a modification of Kripke's theory of truth that would allow classical two-valued valuations. On his account the semantic paradoxes are neither truth-valueless nor both true and false. They are evaluated in the classical way but in a sequence of stages and it turns out that at some stages they are true and at others they are false. Thus they are called unstable in their semantic valuation. The different instabilities yield evaluation patterns and naive semantics exhibits and characterizes them. In 'Notes on Naive Semantics'[41] Herzberger writes:

> What is perhaps unexpected in this reconstruction is the highly systematic character of the semantic instability of paradoxical statements. Far from being incoherent or in any way unmanageable, naive semantics turns out to be no less systematic than Kripke's inductive semantics, and also in some respects stronger than anything accessible within the inductive framework. Consequently, it may afford a sharper characterization of the nature of semantic paradoxes. It certainly does provide an instrument for studying those paradoxes and a new accommodation to them. Rather than attempting to resolve the paradoxes by rendering critical statements truth-valueless or otherwise neutralizing them, naive semantics undertakes to exhibit and characterize their specific patterns and degrees of instability.

Such a characterization of semantic paradoxes is an essential feature of all revision theories of truth. It is a also a distinguishing feature that sets them part from the many traditional approaches trying to 'neutralize' the paradoxes. The revision theories don't try to 'cure' the language of the alleged 'disease' of paradoxes. They consider the paradoxes a natural phenomenon—sometimes comparing them to the eclipses—which requires explanation not expurgation. They differ also from many traditional approaches that provide ad hoc 'solutions' to the problems. The revision theories' perspective deepens our understanding of truth by offering a unifying principle governing truth, from which both the ordinary as well as the extraordinary behaviour of the concept can be understood. Historically the new approach started in the early eighties with some rather technical observations that were made by Herzberger and Gupta independently of each other.

A philosophical foundation was provided in the late eighties only by Anil Gupta,[42] and later by Gupta and Belnap in their book 'The Revision Theory of Truth.' In his book we find the technically as well as philosophically richest revision-theoretic approach up-to-date. Even though some technicalities have been challenged, Gupta's and Belnap's philosophical interpretation is accepted and shared by contemporary revision theorists. It appears thus appropriate to regard their work as the standard revision theory. The present survey will concentrate on the standard theory and the radically new understanding of the concept of truth proposed therein.

The central idea of revision theory is that the concept of truth is defined by a circular definition. Traditionally circular definitions have been ruled illegitimate. Yet, Gupta and Belnap show a way how to use these definitions without incurring the problems traditionally associated with them. In the following section the key ideas of how to make sense of circular definitions are explained. The next section explains the specific application to the concept of truth.

Circular definitions

Gupta and Belnap observe a striking similarity between the behaviour of concepts defined by certain kinds of definitions and the behaviour of the

concept of truth. They show how rules for the introduction and elimination of the definiendum of such definitions parallel the rules for the introduction and elimination of 'true' The definitions they investigate are circular, as for example the following[43]:

$$Gx =_{\text{Def.}} (Fx \wedge Hx) \vee (Fx \wedge \neg Hx \wedge Gx) \vee (\neg Fx \wedge Hx \wedge \neg Gx),$$

which says that the x which are G are either F and H or they are F, non-H and G, or non-F, H and non-G. Suppose we are talking about four objects a, b, c, and d. Suppose further that a and b are F and c and d non-F, and that a and c are H and b and d non-H. Can we determine which are G? The difficulty is that to determine which objects are G we should already know the extension of G. The simple yet ingenious solution to this problem is to assume a hypothetical extension of G and then calculate under this hypothesis which objects are G. This method does not give us an absolute extension of G, instead we get an extension that is relative to the assumed extension of the definiendum. Thus using the definiens to calculate a new extension *revises* the assumed extension, whence the name *revision theory*.

To work out an example in detail should be helpful. Assume thus that nothing is G. Under this assumption we can now evaluate the definiens for each of our four objects. $(Fa \wedge Ha) \vee (Fa \wedge \neg Ha \wedge Ga) \vee (\neg Fa \wedge Ha \wedge \neg Ga)$ is true because the first disjunct is, and hence the definiendum Ga holds. So under the hypothesis that nothing is G the definition rules that a is G. Consider $(Fb \wedge Hb) \vee (Fb \wedge \neg Hb \wedge Gb) \vee (\neg Fb \wedge Hb \wedge \neg Gb)$. The first disjunct is false since b is F but not H. The second disjunct is correct on F and H but not on G, under the hypothesis that nothing is G. The third is false as well and so Gb is ruled to be false. That is, b is ruled not to be G. In $(Fc \wedge Hc) \vee (Fc \wedge \neg Hc \wedge Gc) \vee (\neg Fc \wedge Hc \wedge \neg Gc)$ the first two disjuncts fail because c is not F. Yet, the third is true. The hypothesis that nothing is G makes $\neg Gc$ true. Thus if nothing is G the definition rules that c is. Finally, in $(Fd \wedge Hd) \vee (Fd \wedge \neg Hd \wedge Gd) \vee (\neg Fd \wedge Hd \wedge \neg Gd)$ the first two disjuncts fail because d is not F. The third fails because d is not H. So under the hypothesis that nothing is G the definition rules that d is not G.

We have now calculated the first revision stage starting from the hypothesis that nothing is G. Under this hypothesis the definition rules that a and c are G and that b and d aren't. We may now continue these calculations starting from the verdict just obtained. Repeating this process yields a so-called *revision sequence*. For our example starting with the initial hypothesis as stage zero it looks as follows:

revision stage	0	1	2	3	4	...
is G:	–	a, c	a	a, c	a	...

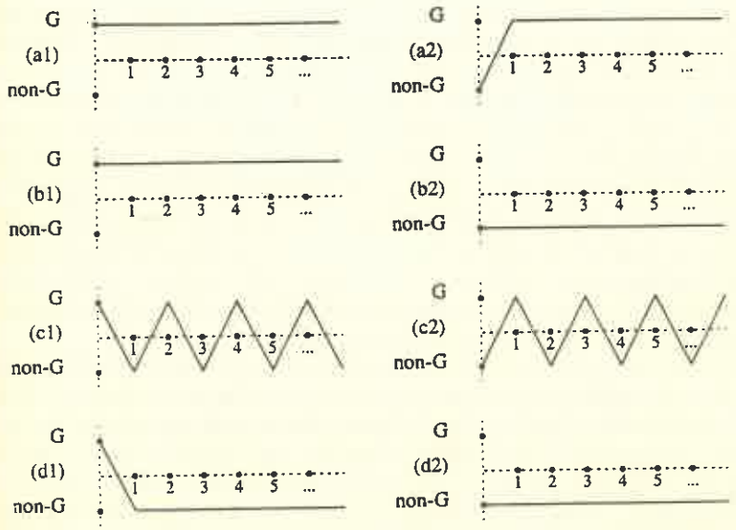
To calculate stage two we go again through the four cases above but now under the hypothesis that a and c are G. For a, b, and d the calculations

will be the same as the ones we just went through. For c the third disjunct of the definiens now fails since $\neg Gc$ is false. It is easy to see that the pattern will continue to have a be G and b and d not be G since nothing changes for these objects. c will continue to switch back and forth since $\neg Gc$ in the third disjunct will make the whole disjunction alternately true when c is assumed not to be G and false otherwise.

In our example we have looked at just one starting point. With four objects there are 16 different starting points possible. If we look at all of them and at all the revision sequences that ensue we find that the four objects behave in very regular ways despite the arbitrariness of the initial hypotheses. Starting each object as G and as non-G we obtain the eight different patterns depicted in Figure 1, where a behaves as in (a1) and (a2), b as in (b1) and (b2), c as in (c1) and (c2), and d as in (d1) and (d2).

The example we worked through starting from the assumption that nothing is G is depicted by (a2), (b2), (c2), and (d2) of Figure 1 taken together. Let us now consider the verdicts for each of the four objects individually.

Figure 1: Revision patterns generated by the definition of *G*



- If the initial hypothesis declares *a* to be *G* it remains *G* throughout. If it declares *a* to be non-*G* the verdict is switched after the first revision stage and then remains. Thus, no matter how we start, eventually *a* is *G* in all possible sequences. If after a certain number of revisions no switching occurs any more, we say a sentence *stabilizes*. If it stabilizes on the same value in all sequences we speak of *convergence*. *Ga* displays both stability and convergence.
- Our calculations showed *b* to be non-*G* in the case of (b2) because all three disjuncts of the definiens failed. Under the hypothesis that *b* is *G* the second disjunct is true, and thus all subsequent stages declare *Gb* true. However, since the stabilities don't agree on one value convergence fails.
- For *c* even stability fails. It keeps switching no matter where we start.
- *d* is very similar to *a*. It is stable in all sequences, and it converges. But it converges on the opposite value, that is *Gd* stabilizes as false in all sequences.

So far we have seen how we arrive at verdicts, albeit hypothetical ones only, by introducing arbitrary initial hypotheses. Now we want to get rid of the arbitrariness and hopefully get categorical judgments. The way to do this is to look at all revision sequences and to check for convergence. Convergence can be observed for *a* and *d*. In these cases we may say that we got rid of the arbitrariness entirely establishing *a* to be categorically *G* and *d* to be categorically non-*G*. We don't have such a verdict for *c*, it remains unstable. The same is true for *b*. Even though *b* is stable in particular sequences, it is not possible to make a categorical judgement when we look at all sequences.

Note also that the process doesn't yield a contradiction as we would obtain in the traditional approach. There we would be forced to conclude that *c* is both *G* and non-*G* which is absurd. To avoid absurdity these kinds of definitions have been ruled illegitimate. Yet, the most exciting outcome of the revision approach is not that we avoid contradictions without having to restrict the the realm of the meaningful. The most exciting feature is that we are able to retrieve good information (with respect to *a* and *d* in the example) even from circular definitions, information that is being missed in a traditional perspective.

To sum up: the general theory of definitions offered by Gupta and Belnap allows all types of definitions, even circular ones. It leaves our ways of working with traditional definitions intact and it expands the realm of the meaningful by showing how we can make sense of circular definitions. We do this in two steps: (1) allow arbitrary hypotheses to get the evaluation started, and (2) get rid of the arbitrariness and arrive back at categorical verdicts (as far as possible) through the revision process. The same stable outcome in all sequences, or 'convergence', is then associated with the unproblematic verdicts. Unstable yet still meaningful patterns emerge in the problematic cases.

For the predicate *G* in the given example, which was chosen to illustrate some of the possible patterns, it is difficult to judge whether what the revision process is yielding is meaningful. We lack intuitions about what ought to be the case with respect to *G*. Not so, however, for the truth predicate. Here we have strong intuitions as to what the verdicts ought to be. And it turns out the revision process delivers precisely these verdicts. It furthermore does so for the unproblematic cases and the problematic alike.[44]

There are other areas where circular definitions can play an illuminating role, testifying to their meaningfulness. Chapter 7 of *The Revision Theory of Truth* discusses *reference, satisfaction, membership, exemplification, necessity, belief,* and *knowledge* as candidates that may be fruitfully reconsidered in this new perspective. The paradoxes of rationality add another vast field for fruitful applications of circular definitions. For this survey of theories of truth, however, it is appropriate to restrict ourselves to the application of the general theory of definitions to the concept of truth.

The definition of 'true'

Recall Tarski' s statement that the T-biconditionals could be considered a partial definition of truth (see section III). This idea encounters tremendous problems because the T-biconditionals can be circular, as in the biconditional '*Tr* (*Tr*(*p*)) iff *Tr*(*p*).' Understood definitionally as $Tr(Tr(p)) =_{def.} Tr(p)$ the term to be defined appears both in the definiendum and the definiens. Thus, if we do not restrict the biconditionals, say by the introduction of levels, the general definition would be circular. This was ruled illegitimate by the traditional theory of definitions. It is only with the general theory of definitions that underlies revision theory that it has become possible to construct this general definition in the simplest and most straightforward way.

Let us represent Tarski biconditionals thus

$$Tr(p_1) =_{def.} A_1$$
$$Tr(p_2) =_{def.} A_2$$
$$Tr(p3) =_{def.} A_3$$
$$\vdots$$

where p_n is a name for A_n. They give rise to the following circular definition of *Tr*.

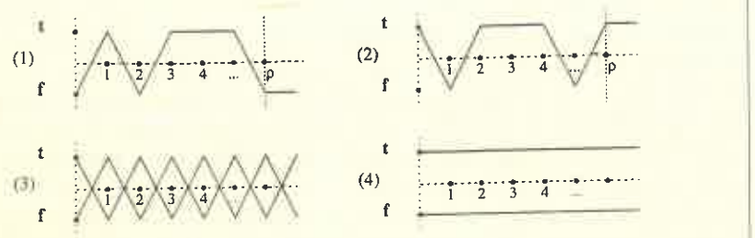
$$Tr(x) =_{Def.} (x=p_1 \wedge A_1) \vee (x= p_2 \wedge A_2) \vee (x=p_3 \wedge A_3) \vee \ldots$$

The circularity of this definition is due to the fact that any of the A_1, A_2, A_3, . . . may contain *Tr*. If we use this (circular) definition to determine the truth of sentences we will find, just as we did above for the predicate *G*, that some keep falling in the extension of *Tr* under repeated revisions,

or they keep falling outside. That is, they *converge* on one of the truth-values. These are the ones that are unproblematically true, or unproblematically false. In the absence of vicious reference all sentences converge in the revision process. No matter what initial hypothesis we start out from we obtain the same and correct verdicts.

With vicious reference present some sentences will behave like the individuals *b* or *c* in the example above, exhibiting some form of instability. Revision theory provides a framework that allows a fine-grained description of these instabilities. It is based on the behaviour of the problematic sentences in all possible revision sequences. Furthermore this description of the unstable cases agrees almost completely with our 'naive' reasoning in these cases. Figure 2 depicts the general case of stabilizing sequences in illustrations (1) and (2).

Figure 2: Unproblematic and problematic cases



The evaluation of a sentence stabilizes (as shorthand we may speak of 'stable sentences') if there is some point ρ after which it doesn't change any more. Up to that point the evaluation may switch between 'true' (**t**) and 'false' (**f**) in any way. Unstable sentences are those for which there is no such point, those that behave erratically for ever. The problematic sentences generate a wide variety of unstable patterns. Illustrations (3) and (4) depict the most common ones, the Liar and the Truth-teller. Both illustrations combine all possible starting points—there are just two in this case—in one diagram. The evaluation in the Liar mirrors precisely the intuitive reasoning we go through when we try to figure out the paradox. If we assume the Liar to be true we must conclude that it is false (is states that it is false, and we assumed that to be true). From its falsity we are led back to its truth (we assume it is false, and it says just that: so it must be true). And the cycle repeats. Assuming that the Truth-teller is true leads to the conclusion that it is indeed true. And continuing to evaluate from there will not change the verdict. Yet, assuming that it is false gives a different verdict. The sentence 'I am true' now incorrectly claims its own truth and must thus be false. Again, this verdict repeats under continued evaluation. Since overall, if we look at all sequences, neither the Liar nor the Truth-teller converges, each of those paradoxes cannot count as categorically true, or as categorically false. It would be incorrect, however, to equate this conclusion with some non-classical logic's third truth-value 'neither true nor false.' Revision theory does not provide a new logic. It has been and is being developed in the framework of classical two-valued logic.[45] So, if the verdict is not 'true,' 'false,' or some other value, what is it? Its 'unstable', or 'oscillating,', or 'paradoxical'. To some it may appear that particularly with the label 'paradoxical' we have come full circle. In a certain sense that is true. Remember that the goal was to explain and not to expurgate certain phenomena. Thus the paradoxes are not supposed to vanish in the process and they don't. In this sense we may be back where we started. Yet, we are back with a deeper understanding of why they will not vanish, why they are supposed to occur. So going in circles was not futile.

Besides shedding light on the origin and the working of the paradoxes, revision theory allows finer distinctions between different paradoxes through a classification of the patterns obtained in the revision process. Such distinctions may seem academic in the truly paradoxical cases. Yet, they become of vital importance in those mixed cases where despite the presence of circularity 'good information' can be extracted. An illustration of such a case is what is known in the literature as 'the Gupta Puzzle'. The main characteristics of this puzzle are that it is circular and yet it has a solution. It constitutes thus a twofold challenge for any theory of truth: the first is not to classify it with the pathological sentences, and the second is to yield the right verdict on it. The Gupta Puzzle has essentially two people, Alfa and Beda, making claims about each other's utterances. They say the following and nothing else:

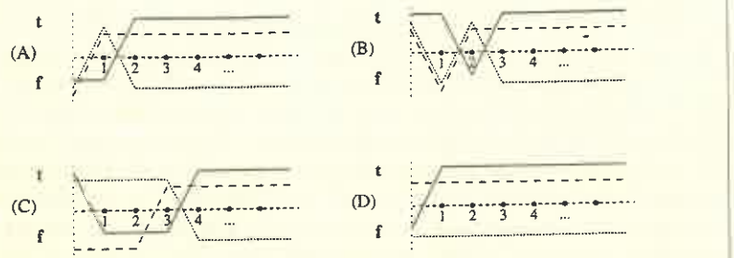
(1) Alfa: everything that Beda says is true.
(2) Alfa: Something that Beda says is false.
(3) Beda: At most one of Alfa's utterances is true.

In a traditional approach we face the problem that the sentences (1)-(3) refer to each other and thus our evaluation does not 'get off the ground'. In revision theory we start with all possible starting points—eight in this case, since we have three sentences that can take two different values—and we observe the stabilities depicted in Figure 3. To save space I have compressed the eight starting points in four diagrams as explained below.

The solid line represents sentence (1), the dashed line sentence (2), and the dotted line sentence (3). The diagrams (A) and (B) start out with the arbitrary assignment '**f, f, f**' and '**t, t, t**'. Had we started the revision with the assignments at the stages 1 or 2 of diagrams (A) and (B) the pattern would be the same as it is from these points on. Thus we have

covered 5 possible starting points. The remaining three are covered in diagrams (C) and (D). What these diagrams mirror is again the intuitive thought process we go through if we imagine, for example, that all three sentences were false (as in diagram (A)), and then try to figure out what would have to be the case. The interesting outcome is the stability from stage 4 on. No matter where we start, it turns out, according to the revision process, that Alfa's first utterance is true, that his second is false, and that Beda's utterance is true. This outcome corresponds to the solution of the puzzle which can be put as follows: Alfa's two utterances couldn't be true simultaneously since they contradict each other. Beda remarks just this and hence his utterance is true. But then everything he says is indeed true, and so Alfa's second utterance must be false.

Figure 3: The Gupta Puzzle



Summary

Starting from the observation of a striking similarity between the behaviour of circular definitions and the concept of truth, Anil Gupta and Nuel Belnap have developed a theory of truth that views the Tarski biconditionals definitionally and that doesn't restrict them to avoid circularity. Instead they consider the circularity a natural feature of the concept and they provide a general theory of definitions based on which the circular definition of truth acquires meaning.The key feature is to understand the concept 'true' as governed by a rule of revision, rather than a rule of application known from ordinary predicates. Rather than dividing the domain to which a predicate, say 'rhombus', applies into two subdomains, the class of the things that are rhombuses, and the class of the ones that aren't, 'true' gives us only a hypothetical extension of the propositions that are true. If there is no vicious reference present in the language the categorical judgements we recover from considering all revision sequences will give us two definite classes of the true and the false

propositions. However, in the presence of pathological propositions, such as the Truth-teller or the Liar, we are left with a different picture. There are still those propositions that are unproblematically true or false, but others may switch back and forth in the process of revision, or they may not be in the extension of 'true' in all sequences. Interestingly the revision patterns we obtain correspond very closely to the intuitive reasoning we perform in these pathological cases. The revision theory of truth is thus able to give us a unified and satisfying account of both the ordinary and the extraordinary features of truth.

VI. THE STRENGTHENED LIAR

By the strengthened Liar I understand the phenomenon that using the terms of a given 'solution' to the paradoxes we can construct a new paradox. A simple example would be the 'solution' that declares the Liar sentence meaningless. The strengthened Liar for this 'solution' is: 'This sentence is either meaningless or false' If we apply the same solution and hold that that sentence is meaningless, then we note that this is exactly what it says in the first disjunct. Therefore it should be true. But if it is true it could not be meaningless. If we assume it to be true then either it is meaningless, in which case it is not true, or it is false, in which case it is also not true. And finally, if we assume that it is false, then both disjuncts are false, that is, the sentence is meaningful and not false. So we arrive at the same contradictions as in the Liar case.

Similarly for 'solutions' that introduce new truth values or allow valuations to be partial. 'This sentence is either gappy or it is false' is a strengthened Liar for a three-valued 'solution' that allows the third truth value 'GAP.' If we assume that it is gappy it would be true, for this is what the first disjunct says, hence not gappy. If it is true it must be either gappy or false, both of which mean that it is not true. And if it is false it must be not gappy and not false.

It has been observed that a similar phenomenon occurs for the revision-theoretic account. 'This sentence is either unstable or it is false' is the strengthened Liar we have to consider. Clearly, if we give a revision-theoretic account of it on which it turns out to be unstable (like a Liar), then its first disjunct is true and it should thus be true and hence stable. If we assume it to be stable, then the first disjunct is false. But then the second will tell us that it is false if we assume it to be stably true and it would appear to be true if it were indeed false. In other words: it would behave like a Liar, that is, it would be unstable.

The response to this challenge will be different for each of the theories presented in this paper. As an example I will present a possible revision theoretic response. Yet, in so far the problems raised are of a general nature the reply may apply for other approaches as well. The first reply is to draw attention to the fact that the goal was not to elucidate all kinds

of pathologicalities but the ones arising from adding 'true' to a simple first-order language for which it was assumed that it contains no further difficulties.[46] To do so, new notions were introduced, such as 'categorical,' 'unstable,' which are not part of the language itself but part of the meta-language used to describe a language containing a truth predicate for itself. It is correct that similar pathologicalities arise if we add the new concepts to the language, which suggests of course a similar explanation, that is, that they are governed by revision rules, thus giving rise to yet other concepts of categoricalness, etc. It is incorrect, however, to judge the project of explaining truth by the success of the much larger project of explaining all new pathologicalities arising from the tools we have introduced to shed light on what interested us first and foremost, namely the concept of truth. A careful and detailed investigation of the different concepts we need to appeal to has not yet been undertaken. Yet, the authors criticizing the revision theory because of this appeal to a Tarskian hierarchy[47] are not impatiently awaiting the results of such an investigation. The very appeal to the notion of a 'hierarchy', so it seems to them, discredits the whole project of explaining truth. Their dissatis-faction is due to the fact that we haven't provided a so-called 'universal language' in which 'everything formulable by language can be expressed in this language.'[48] Or more cautiously: 'perhaps English cannot express every concept that is expressible in some language; but it seems that it can express all of its own semantic concepts'.[49] This is called 'semantic universality' by Simmons which revision theory doesn't provide either.

There are two problems here that should be clearly separated. One is the problem of giving a good account of the concept of truth, and the second is giving a good account of the working of language in general. We cannot and we need not wait for the solution of the second problem before we start to tackle the first—we cannot because we would never start, and we need not if we don't make the tacit assumptions that paradoxes are to be 'solved' all at once or none at all. Thus it is legitimate to offer a picture of the workings of truth without having an all encompassing theory of language, that is, to solve one problem at the time. To illustrate the perceived problems I quote two passages from contemporary authors.

> If a solution to semantic antinomies claims to be a global theory of language in which the difficulties posed by the antinomies are banned, then the meta-theoretic propositions being also formulated with the means of language are *objects of that theory*. So if we can form new antinomies in the sense of the strengthened Liar with the terms of the solution, or if we get counter-intuitive results or results not being in accord with the fundamental ideas of the theory, then such a theory of language must be considered failed.[50]

It is being asked too much of a theory of truth that it should also provide a 'global theory of language' banning all possible paradoxes. Of course, such a theory may be attempted and would certainly be a giant step, but it is hasty to conclude that a theory of truth has failed because it didn't provide 'the' or a general solution. Even a small step in the right direction is an advance and should be appreciated. A related point is made by Simmons.

> Moreover, once Tarskian hierarchies are admitted for even some of our ordinary semantic concepts, then Herzberger and Gupta face the problems that any Tarskian resolution faces. We can argue that in natural language there is just one predicate 'semantically unproblematic sentence' and just one predicate 'semantically unproblematic truth', and it is artificial to split these predicates into infinitely many distinct predicates, each defined with respect to a distinct language.[51]

By 'semantical unproblematic truth' is meant a categorical truth. The complaint is then that we have just one notion of categorical rather than categorical$_i$, for $i \subset$ N, just as with 'true'. Simmons objects to treating the two semantic notions differently, claiming that 'a 'mixed' theory that treats them differently is unfaithful to our ordinary usage'.[52]

Note two problems: (1) The term 'categorical' is not part of our 'ordinary usage.' It was introduced as a technical term into our language by the authors of the revision theory. On the other hand, the term 'true' they want to describe is not a technical term they introduced, but the ordinary everyday notion of truth. Thus there are grounds for treating them differently. There is a case to be made for giving an account of 'true' that doesn't 'split the predicate into infinitely many predicates' precisely because it is not a notion we introduce but one we find in natural language and that we want to describe faithfully. But not so for 'categorical,' it was introduced as a tool, thereby extending our language. This leads to (2): It is highly doubtful that the notion 'one (universal) language' describes any natural language. English, like all natural languages, is transformed constantly, some words and phrases are no longer used, or used with different meanings, and new words and phrases are added whenever the need arise as for example with the term 'categorical'. It seems thus not implausible to claim that English is a universal language in the sense that we can always add resources if the need arises. It does not seem plausible, however, to assume that at any given moment the language does contain all these resources.[53] The appeal to a Tarskian hierarchy of theoretical terms 'categorical$_i$' does in no way conflict with this sense of universality. Finally, the reappearance of paradoxes does not undermine the 'solution' for revision theory as it does for other approaches. Its explanation of the strengthened Liar phenomenon reinforces the 'solution.' If we add 'categorical in L' to the language L we incur the same problems as we had with truth. Think of sentences that

claim of themselves that they are categorical. Their evaluation requires again a hypothesis as to which sentences are categorical. Not surprisingly 'categorical' works very much like 'truth.' This suggests a revision-theoretic explanation which parallels the one for truth. In this perspective the strengthened Liar becomes a natural phenomenon like the Liar, but one that we now understand much better.

In sum: The strengthened Liar in the sense that the terms of the 'solution' lend themselves to the construction of a new paradox poses a problem for all theories of truth. The revision theory is the only one that offers an account of these problems which fits perfectly in its overall perspective.

Those who reject this perspective and appeal to the strengthened Liar to make their case go wrong in at least two ways: (1) The success of a theory of truth which is a theory precisely of the predicate 'true' and not one of the functioning of natural language as a whole is not dependent on the success of the larger project.

(2) The question whether English is a 'universal language,' or even just 'semantically universal' is not as clear as some seem to think. It needs to be clearly specified what 'universal' means. Depending on the answer, hierarchies might well play some role in a general theory of language. Accepting that hierarchies may play a role does not mean that we might just as well have stayed with Tarski's account of truth. On the other hand, being dissatisfied with Tarski's account does not mean that we should reject all hierarchies. They may be satisfactory for some purposes but not for others.

FINAL REMARK

Paradoxes and pathologicalities have played an important role in the development of the many theories of truth available today. Some theories have been motivated by strong intuitions of their authors about particular paradoxes. All theories make an attempt to integrate an explanation of the ordinary and the extraordinary behaviour of the concept of truth. How convincing and how illuminating an approach is will be determined by the correct weight given to the different intuitions in this explanation. And it will be determined by how natural the accommodation of these intuitions in the theoretical framework appears. The bewildering richness of just the theories that have flourished over the past thirty years does not make it easy to judge which theory fares best on this count. But I hope that the general road map provided in this survey will facilitate the readers' own inquiry into the details of the recent theories of truth and after that let them be the judge.

NOTES AND REFERENCES

1. Aristotle, *Metaphysics*, 1011b, 23.
2. Although the paradox is sometimes named after Epimenides there is no evidence that he noticed the problem. It is Eubulides of Miletus, a contemporary of Aristotle, who is credited by Diogenes Laertius with the discovery of the paradox. For more details on the early history of the Liar paradox see Alexander Rüstow, 'Der Lügner: Theorie, Geschichte und Ausflösung', PhD thesis, Universität Erlangen, Erlangen, Teubner, Leipzig, 1910.
3. A particularly striking example, the so-called Gupta-Puzzle, is discussed in detail in section V.
4. Alfred Tarski, 'The semantic conception of truth', *Philosophy and Phenomenological Research*, Vol. 4, 1994, pp. 341–76, cf. p. 344.
5. Alfred Tarski, op. cit., p. 20.
6. Anil Gupta, 'Truth and Paradox', *Journal of Philosophical Logic*, Vol. 11, 1982, pp. 1–60; reprinted in *Recent Essays on Truth and the Liar Paradox*, R.L. Martin (ed.), Oxford University Press, Oxford, U.K., 1984, pp. 175–235 (page references are to this edition).
7. They may be found in any contemporary introduction to formal logic.
8. Alfred Tarski, 'Der Wahrheitbegriff in den formalisierten Sprachen', *Studia Philosophica*, Vol. 1, 1935, pp. 261–405; English translation by J.H. Woodger in 'The Concept of Truth in Formalized Languages', in A. Tarski: *Logic, Semantics, Meta-Mathematics*, Oxford University Press 1956, 2nd edition published by Hackett, Indianapolis, 1983, pp. 152–278, cf. p. 165 and p. 267.
9. Tyler Burge, 'Semantical Paradox', *Journal of Philosophy*, Vol. 76, 1979, pp. 169–98, reprinted in *Recent Essays on Truth and the Liar Paradox*, R.L. Martin (ed.), Oxford University Press, Oxford, U.K., 1984, pp. 83–117, cf. p. 94.
10. Ibid.
11. Ibid., p. 105.
12. Cf. axiom (6'): $\neg \mathrm{R}i(\Gamma, \alpha) \supset \neg \mathrm{Sat}i(\Gamma, \alpha)$, see Tyler Burge, 'Semantical Paradox', op. cit., p. 105.
13. Ibid., p. 100.
14. Tyler Burge, ' The Liar Paradox: Tangles and Chains', *Philosophical Studies*, Vol. 41, 1982, pp. 353–66.
15. Tyler Burge, 'Semantical Paradox', op. cit.
16. Tyler Burge, 'Semantical Paradox', op. cit.; 'The Liar Paradox: Tangles and Chains', op. cit.
17. Haim Gaifmann, 'Pointers to Truth', *Journal of Philosophy*, Vol. 89, 1992, pp. 223–61.
18. Haim Gaifman, 'Operational Pointer Semantics: Solution to Self-Referential Puzzles I', *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge held at Los Altos, California*, Moshe Y. Vardi (ed.), Morgan Kaufmann, 1988, pp. 43–59.
19. Haim Gaifmann, 'Pointers to Truth', op. cit.
20. Ibid., p. 242.
21. Ibid., p. 244.
22. Ibid., p. 243.
23. Ibid.
24. Using the more familiar N ('neither') for GAP the strong Kleene tables for negation ($\neg$) and conjunction ($\wedge$) are as follows:

$\neg$	
T	F
F	T
N	N

$\wedge$	T	F	N
T	T	F	N
F	F	F	F
N	N	F	N

The combinations of the values T and F are evaluated in the classical way. Classically the negation of a true sentence is false, and vice versa. The negation of a neither true nor false sentence, a 'gappy' sentence, is also gappy or N. Conjunctions are true if and only if both conjuncts are T. And conjunctions are false if and only if one conjunct is F. This holds classically and it is preserved by the strong Kleene tables for cases involving N. That leaves the combinations T/N and N/N which are both defined to be N.

25. Haim Gaifman, 'Pointers to Truth', op. cit., p. 249.
26. Peter Aczel, *Non-well-founded Sets*, CSLI Publications, Stanford, 1988.
27. John L. Austin, 'Truth', *Proceedings of the Aristotelian Society*, supplementary volume 25, 1950, reprinted in *Philosophical Papers*, J.O. Urmson and G.J. Warnock(ed.), Oxford University Press, Oxford, U.K., 1961, p. 122.
28. Liar cycles, as the one Plato and Aristotle participate in, may also have an asymmetry between the participants. That is, one might speak truly and the other falsely—it depends on the situations their propositions are about. Cf. Jon Barwise and John Etchemendy, *The Liar: An Essay on Truth and Circularity*, Oxford University Press, Oxford, U.K., 1987, p. 149.
29. This might be considered the main source of ambiguity. The authors also mentioned ambiguity due to a conflation of denial and negation, and ambiguity introduced by the fact that the English 'this' (which is their device to achieve self-reference) has a demonstrative and a reflexive use.
30. Gupta makes this point in 'Jon Barwise and John Etchemendy, The Liar: An Essay on Truth and Circularity', *Philosophy of Science*, Vol. 56, 1989, pp. 697–709.
31. Robert C. Koons, Paradoxes of belief and strategic rationality, Cambridge University Press, Cambridge, U.K., 1992, pp. 103–04.
32. Saul Kripke, 'Outline of a Theory of Truth', *Journal of Philosophy*, Vol. 72, 1975, reprinted in *Recent Essays on Truth and the Liar Paradox*, R.L. Martin (ed.), Oxford University Press, Oxford, U.K., 1984, pp. 53–81, cf. pp. 64–65.
33. Ibid., p. 74.
34. Ibid., p. 63.
35. Going beyond classical logic are the four-valued approaches in Peter W. Woodruff, 'Paradox, Truth and Logic', *Journal of Philosophical Logic*, Vol. 13, 1984, pp. 213–32; Albert Visser, 'Four-valued semantics and the Liar', *Journal of Philosophical Logic*, Vol. 13, 1984, pp. 181–212; Graham Priest, 'The logic of paradox', and 'Logic of paradox revisited', *Journal of Philosophical Logic*, Vol. 8, 1979, pp. 219–41; and Brian Skyrms, 'Intensional Aspects of Semantical Self-reference', *Recent Essays on Truth and the Liar Paradox*, Robert L. Martin (ed.), Oxford University Press, Oxford, U.K., 1984, pp. 119–31. Skyrms' approach is described in section IV.
36. Brian Skyrms, Return of the Liar: Three-valued Logic and the Concept of Truth', *American Philosophical Quarterly*, Vol. 7, 1970, pp. 153–61.
37. Brian Skyrms, 'Intensional Aspects of Semantical Self-reference', op. cit.
38. Ibid., p. 126.
39. Ibid., pp. 124–25.
40. Hans Herzberger, 'Notes on Naive Semantics', *Journal of Philosophical Logic*, Vol. 11, 1982, pp. 61–102, reprinted in *Recent Essays on Truth and the Liar Paradox*, op. cit., pp. 133–74 (page references are to this edition).
41. Ibid., pp. 134–35.
42. Anil Gupta, 'Truth and Paradox', op. cit; and 'The meaning of truth', op. cit.
43. This example is treated in more detail in Chapter 4 of *The Revision Theory of Truth*, Anil Gupta and Nuel Belnap, The MIT Press, Cambridge, MA, 1993.
44. I do not want to claim strong intuitions for *all* cases involving truth. They do get weaker in some pathological cases. This has led to some disagreement over the 'correct' verdicts in these cases, and consequently to technical alternatives which

'fine-tune' the revision process differently. An example is *The Liar Speaks the Truth: A Defense of the Revision Theory of Truth*, Alladin Yaqūb, Oxford University Press, Oxford, U.K., 1993. Discussions of technical alternatives may be found in Nuel Belnap, 'Gupta's Rule of Revision Theory of Truth', *Journal of Philosophical Logic*, Vol. 11, 1982, pp. 103–16; and André Chapuis, 'Alternative Revision Theories of Truth', *Journal of Philosophical Logic*, Vol. 25, 1996, pp. 399–423.

45. This is not a requirement, however. The approach applies equally well to non-classical frameworks. The view among revision theorists is that even though truth is a fundamental concept it should not decide on a particular logic but work with all of them.
46. This simplifying assumption can itself be challenged of course (Van McGee does so in *Truth, Vagueness and Paradox: An Essay on the Logic of Truth*, Hackett, Indianapolis, 1991, pp. 146–47), but a solution that introduces more complications should give plausible reasons independent of the paradoxes for such a complication. I cannot do justice in a few sentences to those authors that disagree because of this, the issue would be the different fundamental intuitions about truth and about logic that we want to preserve by a theory of truth, and it would be the question which intuitions we are willing to give up. Suffice it to say that the revision theory requires very little in this respect. It is built exclusively on the Tarski biconditionals which is the most deeply held intuition by all theoreticians of truth, and it does not even restrict them.
47. Cf., for example, Elke Brendel, *Die Wahrheit über den Lügner: eine philosophisch-logische Analyse der Antinomie des Lügners*, de Gruyter, Berlin, New York, 1992, pp. 53–54; Van McGee, 1991, op. cit., p. 147; and Keith Simmons, *Universality and the Liar: An Essay on Truth and the Diagonal Argument*, Cambridge University Press, Cambridge, U.K., 1993, pp. 67–69.
48. My translation of Elke Brendel's *Die Wahrheit über den Lügner: eine philosophisch-logische Analyse der Antinomie des Lügner's*, op. cit., p. 133: 'eine[r] universelle[n] Sprache, in der alles sprachlich Formulierbare auch innerhalb dieser Sprache ausdruckbar ist'.
49. Keith Simmons, *Universality and the Liar: An Essay on Truth and the Diagonal Argument*, Cambridge University Press, Cambridge, U.K., 1993, p. 15.
50. This is my translation of the following quote from the original German: 'Erhebt ein Lösungansatz für semantische Antinomien den Anspruch, eine globale Theorie der Sprache zu sein, in der die Antinomieproblematik gebannt ist, so sind metatheoretische Aussagen, die ja ebenfalls mit sprachlichen Mitteln formuliert sind, *Gegenstände dieser Theorie*. Wenn sich somit im Sinne des verstärkten Lügners mittels der Terminologie dieser Lösungsansätze erneut Antinomien bilden, bzw. kontraintuitive oder mit den Grundideen der Theorie nicht in Einklang stehende Resultate erzielen lassen, so muss eine solche sprachliche Theorie als gescheitert angesehen werden', Cf., Elke Brendel, *Die Wahrheit über den Lügner*, op. cit., pp. 53–54.
51. Keith Simmons, *Universality and the Liar*, op. cit., p. 69.
52. Ibid., fn. 13, p. 196.
53. Simmons' own solution (the 'singularity theory') ends with the following observation: 'We can say everything there is to say, but not all at once', ibid., p. 182] which would seem to capture the same general insight.

Rajendra Prasad—'Some Comments on Applying Ethics'

The second section ('Applicability of an ethical theory') of Dr Rajendra Prasad's interesting and clearly-written essay, 'Applying Ethics' (*JICPR*, Vol. XIV, No. 2, 1997) has provoked the following comments.

An 'ethical theory', according to Dr Prasad, tells us, 'in a systematic, conceptually organized way', 'what sorts of actions we should do', 'what could be a moral duty', etc. And he tells us that 'solving an ethical problem... involves applying an ethical principle, or point of view, to the situation concerned' (pp. 5–6). I believe that in enlarging on this thesis Dr Prasad gives a rather unrealistic picture of moral activity and one which seems incompatible with his own further views.

It is not clear what Dr Prasad exactly means by an ethical theory. He could not be meaning meta-ethical theories like Emotivism or Prescriptivism (though he writes {p. 4} that he is here not distinguishing between normative ethical and meta-ethical theories), because such theories are about the nature and status of moral judgements and cannot give any specific moral guidance. Does he mean theories like Hedonism, the Categorical Imperative (which is partly meta-ethical) or Self-realization, Consequentialism or Divine Command theories? The last he himself dismisses as not being able to tell us what we should do, and the same sort of fault could be found with Consequentialism, for it cannot tell us what, in a particular case, our personal or social well-being might be or, indeed, what 'well-being' might mean. Similar remarks might be made about the other theories: they cannot tell us how to realize ourselves or what our duties are or how we could tot up units of pleasure even if we knew what exactly is to count for a pleasure.

Dr Prasad talks of 'norms, principles, criteria of rightness' (p. 10) and it is these that he probably means by 'ethical theories'—something like the Ten Commandments or good parental advice or copy-book maxims like 'Never a borrower or a lender be' or 'Return what is not yours' or 'Be honest, come what may'. To every one of these and dozens of similar others, exceptions can be found. It is not always right, for example (as Plato shows in *Republic*, book 1), to return what one has borrowed. Circumstances alter cases and the exceptions cannot possibly be exhaustively enumerated or even thought of at any one time. Dr Prasad admits as much. In a 'tricky situation', he writes, one 'may fail to see how the relevant theory or theories are to be applied', for the problem may be 'unusually complicated' and human intelligence is finite, so that 'the

application of ethical knowledge has become so difficult' (p. 5); one may then 'have to decide with how much rigour. . . the relevant theory is to be applied in order to have a fit between the situation and the theory' (p. 7); one may have to 'explore which ethical theory can be relevantly applied' (p. 7). The question is: in order to do all this, does one apply a super-theory or a super-principle? Dr Prasad seems to think that, in tricky cases, one needs to consider the 'technical, sophisticated, precizified version' of the theory (p. 5). To do this is admittedly beyond the capacity of ordinary people, but even a Bhīṣma expresses his inability to apply a theory, not because of the finitude of human intelligence (as Dr Prasad earlier suggests), but because the theory fails to 'clearly specify the conditions' which must be satisfied for its proper application (p. 9) or, to put it differently, the theory does not specify (and cannot specify) the innumerable exceptions to its proper application. And even if a theory did specify these conditions and exceptions, Dr Prasad himself adds, one may still 'question their validity' (p. 9). So what remains of the guiding function of an ethical theory?

In the face of such difficult situations, Dr Prasad seems, on the one hand, to hold that one has to take the help of 'experts' (p. 10), 'acknowledged possessors of ethical knowledge' (p. 5), though even they, like Bhīṣma, may find a theory so 'extremely subtle and deep' that they would not know how to apply it. The idea of moral experts was long ago proposed by Plato, who held that, once such experts were appointed Guardians of a state, the citizens would simply have to accept whatever the former, in their absolute wisdom, might decide to be good for the latter. Not only has this idea been largely discredited, but the whole point of moral action is that you do not simply follow anyone's command or even advice—not even God's—but must come to your own decision after considering the situation from the widest possible angle and according to your moral insights, such as they might be. You may seek advice but you need not have to follow it. Dr Prasad's analogy about a psychological theory being 'usable in making us better learners' (p. 6) does not apply here, for such a theory only recommends the means for achieving already decided ends. In the case of moral action, we are concerned with ends not yet decided or, more correctly, the ends-means idea does not strictly apply. The 'what' and the 'how' are inextricably interwoven.

On the other hand, Dr Prasad holds that an ethical theory, even if formulated by a 'highly gifted individual' of 'elevated character', is 'meant for normal people who have normal abilities'; 'guidance. . . is needed more by ordinary than by extraordinary people' (10). He says that 'a general understanding of the broad aspects of an ethical theory is enough for applying in normal situations' and that 'in the majority of cases we know how to apply the available ethical knowledge' (p. 5).

I submit that this entire idea of *applying* moral knowledge or moral principles is misconceived. In normal cases people do, of course, act morally, but they do not do so by applying available ethical knowledge. They act naturally and spontaneously. A man tells the truth because there is no reason to tell a lie; one returns a borrowed article because one just does not think of appropriating it. Recalling a relevant principle, applying it, deciding to act on it and then acting on it—all this just does not happen. One's action is simply, in many cases, the expression of one's mature, cultivated moral sense, 'partly inherited', as Dr Prasad so well puts it, from one's 'cultural heritage and partly acquired through his own reflection on certain situations' (p. 7).

In such everyday situations of moral life, there is no moral conflict and no application of principles. Indeed, in many cases, if you ask the person concerned what principles he applied, he may look at you in wonderment and not be able to answer. Only when a moral *problem* arises does one have to pause, reflect and decide to act in a certain way. A man's mother is dangerously ill, only an imported drug can save her, but the customs officer refuses to release it unless 'the usual' is done. Suppose the person concerned does not believe in doing 'the usual' and has never done it so far. Here is a 'tricky' case. One principle is fairly clear: don't bribe a public servant. What is the other? In such a situation, is one aware of a principle, leave alone formulating it? Does one think of the principle of not bribing and then go through all the possible exceptions to it? Are the exceptions based on any principle? In any case, a conflict of principles is not resolved by appealing to a superior principle. Perhaps, given enough time, one may be able to think out such a superior principle, but certainly not at the time of acting. The principle could perhaps be worked out by reflecting on the act; the act is not performed by applying the principle.

Finally, Dr Prasad says, quite correctly, that 'newer and newer ethical problems' keep arising because 'the human world and human nature are extremely complex and non-static' (p. 11). What principles would one then apply when moral problems arise? Would you formulate a new principle and then apply it? Or would you first assess the moral problem and solve it according to your moral insight and then, perhaps, formulate a suitable principle? 'Sometimes', writes Dr Prasad, 'it may be difficult to make a choice of the relevant principle' and 'on the theoretical level it may not be possible to say with full confidence which theory is best', but he assures us that 'in practical life the situation is not so bad', for 'the demands of life almost always show the way if one has the will' (p. 11). Even if one is inclined to accept this rather optimistic picture, it still does not mean that first you have a principle and then you apply it.

121 Pushpak, Altamount Road, Mumbai S.K. Ookerjee

Prabhakara Rao on 'Brahman-realization'

While it is admirable that Prabhakara Rao has tried to raise what appear to be new questions and issues, the clarity with which he has attempted this task leaves much to be desired.[1] He speaks of 'personal experience' (p. 71), 'experiential proof' (p. 72) and 'practical realization' (p. 72) in terms of which he demands Brahman-realization to be made evident, but nowhere in his article does he tell us what those expressions mean. Particularly because all Advaitins claim their discourse about Brahman to have a firm experiential foundation, it was all the more necessary that he should have been very clear about the meaning of these expressions according to him.

The central notion of 'hypothesis' which he employs throughout his paper is itself confusing and unclear. He says that 'This Ātman is Brahman' (p. 72) 'is basically conceived as a hypothesis in Advaita Vedānta' which, according to him, is 'based on three important premises' and one of these premises is 'This Ātman is Brahman' itself! It is not at all clear whether or not there is any difference between 'hypotheses' and 'premises' according to him and this makes it very difficult to follow what he is saying or arguing for.

Many questions arise because clear definitions are not offered. What about the 'personal experiences' of a Sri Ramakrishna or a Ramana Maharshi? Does Prabhakara Rao accept the possibility of some kind of 'personal experience' in their cases or does he believe that these men were deluded, or were faking? If Prabhakara Rao rejects the claims of Brahman-realization made by them or on their behalf by others, does he do so on the ground (a) that their experiences have no authority or relevance for him (thus implicitly conceding the possibility of their having had such experiences and thereby also conceding the possibility of such experiences) or, on the ground (b) that they could not have had such 'personal experiences' (because there are no such experiences at all)? I can readily concede ground (a). But in the event of ground (b) being relied upon by Prabhakara Rao I must point out that in asserting that they could not have had such experiences he is not on any ground stronger than those who assert that they had such experiences. Or, does he mean by "personal experience" his own experience? In that case, should we accept that there is nothing like Brahman-realization because Prabhakara Rao did not have it? Or, again, when he demands to know whether Brahman-realization is borne out by "personal experience," is he using the Positivist notion of "empirical experience" to reject *Brahmānubhava*? If he has relied on this well articulated and defined notion of empirical experience and the associated idea of verification, it is better that it be declared so in clear and unambiguous terms.

If Prabhakara Rao is using the long-discarded Logical Positivist theory of verification with its associated notion of empirical experience, I wish to point out very briefly that the form in which he would be required to use that theory to dismiss metaphysics of the Advaita Vedāntic type concerning Brahman-realization would also automatically lead to the dismissal of virtually the whole of scientific theory of the twentieth century.

Prabhakara Rao also makes several statements for which it is hard to find justification. For example, I do not know which is the Indian philosophical school which is a 'pursuit that disagrees with both belief and disbelief in the existence of self' or, for that matter, even the 'pursuit based on the disbelief in the existence of the self' (p. 71). Also, he says that 'most of the Upaniṣads begin with an inquiry into the Reality from the unknown Brahman to the known world . . .' (p. 71), but is it true? If anything, the inquiry in the Upaniṣads as a rule proceeds from some general question concerning the known world (like '*Kasminnu bhgavo vijñāte sarvāmidaṃ vijñātaṃ bhavati. . .*' in *Muṇḍakopaniṣad*, '*Keneśitaṃ patati preṣitam manaḥ kena prāṇaḥ. . .*" in *Kenopaniṣad*, '*Astītyeke nāyamastīti caike . . .*' in *Kaṭhopaniṣad*, and '*Kuto ha vā imāḥ prajāḥ, prajāyanta. . .*' in *Praśnopaniṣad*, etc.) and it ends up with some thesis about Brahman.

The 'unknown Brahman' is not the starting point of inquiry of the Upaniṣads and the inquiry does not usually proceed from Brahman to the world but mostly in the opposite direction. But why does Prabhakara Rao choose not to see it this way? Is it because he is already persuaded that the 'unknown' Brahman is a 'hypothesis' and that it remains a hypothesis even at the end of inquiry? For him, the 'unknown' also appears necessarily to be the 'unknowable' which is why he does not see the concept of Brahman as the (right or wrong) *logical conclusion of an inquiry* but merely as the 'unproved hypothesis' (p. 72) lying at the beginning as well as the end of an inquiry. That is also the reason why he concludes that if Brahman is the object of faith, it should necessarily remain an object of faith and can never become an object of discovery.

There are also further assumptions for which it is equally hard to find justification. He assumes that (a) if the Advaitin's way of arguing for the existence of *Brahmānubhava* is wrong, faulty, inconsistent or does not make sense, it proves that there is no such thing as *Brahmānubhava*, (b) it is possible to prove that something [like *Brahmānubhava*] does not exist, (c) if anything is 'knowable' it should necessarily also be 'knowable empirically', (d) the absence of cognition of an object is the same as the cognition of the absence of an object, (e) if the knower of the different states of consciousness is the same, then the states are also the same in the sense that if the object cognized by a knower in one state is false, the objects cognized in all other states by the same knower are all false, (f) if something is real, then it should also be experienced, (g) when *pramāṇagata sandeha*, *prameyagata sandeha* and *viparītabhāvanā* concerning Brahman are completely removed, there can *still* be *anubhavagata sandeha* concerning Brahman, and so on.

I wish to make it clear here that I do not at all hold that everything every Advaitin has said is right or reasonable. I feel as much concerned by the laxity with which Advaita is expounded by its ardent followers as by the looseness with which it is attacked by its critics. The most pervasive practice in the world of Advaitic scholarship is the uncritical attribution to Śaṅkara of ideas and viewpoints that he did not hold or could not even conceivably hold. For example, both the followers and critics of Śaṅkara regard the available commentary on the *Māṇḍūkyakārikā* as an authentic work of Śaṅkara even when that work vigorously advocates the theory of absolute non-origination (*ajātivāda*) of the world from Brahman in complete opposition to the equally vigorous exposition of the origination of the universe from brahman in the *Brahmasūtrabhāṣya* by the same Śaṅkara. While the *Māṇḍūkyakārikābhāṣya* attributed to Śaṅkara holds that waking and dream are equal and unreal, is not this idea of their equality and joint unreality systematically rejected in the *Brahmasūtrabhāṣya* of the same Śaṅkara? Did Śaṅkara then freely hold and defend patently self-contradictory theses in different works written at different times? If so, does he offer grounds to reject what he had accepted earlier or, does he just merrily go on expounding a view totally opposed to his own earlier view? Without paying attention to such critical issues, Prabhakara Rao quietly uses the thesis of the unreality of the world found in the *Māṇḍūkyakārikābhāṣya* attributed to Śaṅkara to support his own thesis regarding the impossibility of Brahman-realization while we find Śaṅkara saying in his *Brahmasūtrabhāṣya:* 'Just as Brahman, the cause, never deviates from reality at all the three points of time [past, present and future], the world too, the effect, never deviates from reality at all the three points of time' (*Yathā ca kāraṇam brahma triṣu kāleṣu satvaṃ na vyabhicarati, evaṃ kāryamapi jagat triṣu kāleṣu satvaṃ na vyabhicarati, brahmasūtrabhāṣya*, 2.1.16, p. 160).

The central thesis of Prabhakara Rao as I perceive it is that anything like Brahman-realization is 'not possible'. But, unfortunately, he does not also clarify what he means by this term. Prabhakara Rao seems to be claiming both that (a) 'anything like *Brahmānubhava* is not possible' and (b) '*Brahmānubhava* does not exist ('is false').' Logically, these two claims are different. Claim (a) concerns possibilities and is hence a logical claim and it should be examined purely on formal logical grounds. Claim (b) is a claim concerning existence and has to be settled on non-logical or empirical grounds, if it can be settled at all that way. I certainly concede that non-existence can be proved, but this is possible only when such non-existence is the same as the logical impossibility of existence. We can certainly prove that there are no square-circles because nothing can be both a 'square' and 'not a square' (which a circle is) at the same time. But can we similarly *prove*, for example, that there are no flying horses? This point is quite important and crucial in the context of the whole paper of Prabhakara Rao.

Claim (a) which is purely logical in nature cannot be settled by any appeal to any *facts*. Nor can it be made strong and irrefutable by any amount of *argumentation* for the following reason: it can only be that either Brahman-realization is logically impossible, or is non-logically, contingently or empirically impossible. It cannot be argued that it is logically impossible since only a self-contradiction is logically impossible and *Brahmānubhava* is not a self-contradictory concept. Once it is granted that something like *Brahmānubhava* is logically possible, the only alternative left is to argue that it is empirically impossible. It is impossible to argue that something is empirically impossible by advancing *logical* grounds. Any argument for empirical impossibility has to be grounded in the facts of experience. In that case, while arguing against Brahman-realization Prabhakara Rao is on very weak grounds because, to prove his point, he has to show that *no one in fact can have Brahmānubhava*. But there are works whose reasonable interpretation requires the assumption that those who composed them *had Brahmānubhava*. There have also been hundreds of people on whose behalf *Brahmānubhava* has been claimed, the latest serious case being that of Sri Ramana Maharshi.

It is of course possible that these claims are spurious or fake, or to use his own mild and respectful term, 'false'. This necessarily leads to the question: 'How can anyone know that an experience is true?' Before asking the question 'How can anyone know that *Brahmānubhava* is true?' we shall examine the case of some statement about an empirical experience about which we say we know that it is true. This is necessary because if and only if some statement concerning empirical experience can be shown to be unquestionably true, it can be argued that no statement concerning *Brahmānubhava* is that kind of a statement and therefore it is not true.

Let us take the proposition about an empirical experience embodied in the statement 'The lotus is blue' which is true. It can certainly be granted that there is some kind of an experience which makes two individuals X and Yutter the same statement 'The lotus is blue'. But how do we know that it is the *same* or the *same kind of experience* that makes them utter the same sentence? If, for example, X always experiences blue when there is green and green when there is blue because of some genetic defect in the rods and cones of his eye, growing up in a community of speakers, he will still always use the word 'blue' *correctly* like Y who experiences blue as blue. Here the so-called 'correct description' of a colour by Y is indistinguishable from a systematic '*mis*description' of it by X. Our only available clue to what X is experiencing is his usage of words and if his usage is systematically in agreement with our usage, we conclude that he is having the same experiences as we are having. If another person is using language exactly in the same way we are using, we assume that he is also having the same experiences we are having, but this need not be true at all.

Therefore we cannot appeal to our 'experiences' to certify the truth

of "The lotus is blue" but only to a common usage of the term 'blue'. If everyone is using the term "blue" to describe the lotus, then we conclude that (a) the lotus *is* blue, and (b) that everyone's experience of this blue is the *same*. We really never know what X is experiencing when he uses the term 'blue' and even much less whether that experience of X is the same as the experience we are having when using the term 'blue'. It is one thing to *know* that it is the same (if that can be known at all!), but entirely another matter to *think* that it is the same. We may very well think that it is the same when it is not in fact the same. We can never *know* that our experiences are the same, but that does not at all prevent us from *thinking* that they are the same and even evolving a pattern of successful behaviour on the basis of such thinking.

Thus, it is not at all possible to know that 'The lotus is blue' is true in the sense that we can absolutely validly assert that it is an accurate description of a 'common empirical experience' we all have as human beings. What is true of the statement 'The lotus is blue' is true of *all* empirical statements of any kind. Therefore the commonness of ordinary human experience which is the foundation of all forms of empiricism is simply a belief or a 'hypothesis' and it should always remain a hypothesis. It can never graduate to the level of a fact. Therefore, if there is no single statement involving empirical experience that can be shown to be an absolute truth, i.e., to be a statement involving the same kind of experience as its undeniable foundation, there is simply no basis on which one can proceed to deny the truth of any statement concerning *Brahmānubhava*.

Now, for the sake of argument, let me grant the truth of empirical statements made by several people on the ground that they are inter-subjectively established and that there is a 'common experience' underlying them all. Even then, the possible truth of statements concerning *Brahmānubhava* cannot be denied. This is because every possible form of human experience is not already given, and therefore is not already known, to any one or all of us. There might be forms of experience not already had by any human being so far or there might be experiences that only a few humans have had throughout human history.

More generally, it is impossible to establish conclusively the non-existence of anything which is not logically impossible. Flying horses are not logically impossible objects and if we tour the entire universe and fail to find any flying horses, all that it proves is that *we did not find any* flying horses and not that *there are not any* because a cute, little flying horse might very well get born in the backyard of Prabhakara Rao's bungalow soon after we leave it embarking upon a long journey of the universe to prove that there are no flying horses. Also, as long as something is not a self-contradictory entity, even if all the available arguments for its existence can be conclusively shown to be faulty, inadequate or even logically unsound, it is inconsequential because the impossibility of the existence of such an entity is not at all legitimately a matter of argument. While its

experience can very well be proof enough of its existence, the absence of such experience can never be proof enough of its non-existence.

Note and Reference

1. M. Prabhakara Rao, A Critique on Brahman-Realization, *Journal of Indian Council of Philosophical Research*, January–April, 1977.

Bangalore University, Bangalore Srinivasa Rao

Does Ranjan Umapathy Play Sancho Panza to Paul Churchland's Don Quixote?

In a recent contribution, Ranjan Umapathy[1] expresses major reservations about many theories of mind-body identity that have been operative for the last few decades and proceeds to offer a oneline Upaniṣadistic (cf. *Mandukya Upaniṣad*) refutation of the identity theory (at least in one version) which takes mind as causally dependent on brain. For *Mandukya Upaniṣad*, mind or consciousness is independent and identifiable only with its own existence (Existence Itself) just as dreaming self is dreaming existence: Dreaming is Existence Itself. So the unity-in-difference argument:

> Dreaming is an existence over and above mind or brain
> So, consciousness is an existence over and above mind or brain.

goes through once it is assumed that consciousness is self-existent. This is assumed on pain of the risk of dualism, and once it is supplemented by an alternative proposal (AP) to the erstwhile identity of mental states and brain states, then the so-called identity is blown to pieces. There is unity in lieu of identity. Ranjan treads where angels fear to do so because even Searle who tries to prove such an identity has to meet a similar fate at the hands of Śaṁkara who can refute him by saying that physical process cannot itself become an object of the process. Identity theorists try hard to prove that consciousness is nothing but matter. In brief, they try to argue that if consciousness were reduced to matter, then we can very well know about it. It is a matter of pride to Indian philosophical traditions that if that matter were to be identified with consciousness, then it will know about itself. It cannot know about itself, and so by *modus tollens*, identity does not work. One wonders whether he is deceived from believing that the author is begging a very crucial question, namely that

consciousness can know itself (the other half of Śaṁkara's principle is given as: matter cannot know itself). The following are therefore the motions:

1. 'my mental processes can become objects for themselves' (hypothesis) (p. 42);
2. Śaṁkara's principle: physical processes cannot become an object of the process;
3. (2) contradicts (1);
4. There is something physical which cannot become the object of consciousness;
5. Matter cannot become therefore an object of consciousness;
6. *Ex hypothese,* consciousness is being consciousness of (consciousness has phenomenal character;
7. Consciousness is self-existent.

The above steps prove that consciousness is independent. The reader is then relieved to know that it runs no risk. For once, if they are shown to have an underlying unity, the argument is complete. This leads Ranjan further on to his AP (alternative proposal), which completes the argument (*Infra*). So, the above does not prove that there is an underlying unity. In other words Ranjan invokes this only to wonder: why this (Śaṁkara's) principle cannot be extended to attack all forms of identity theory. Ranjan proposes this as an acid test against Paul Churchland's proposal that I can observe my neural processes provided that our senses are sensitive enough. Churchland suggestion is quixotic because even identity theorists like Shaffer made a suggestion that such a distinguishing feature cannot be identified with brain states. Granting that I can now only feel neurons blasting away in my brain in a very vague way, the onus is on Churchland to explain this only with reference to an appearance-reality distinction, a trademark of Indian philosophical traditions. While Searle appears to recognize this ('Where appearance is concerned we cannot make the appearance-reality distinction because the appearance is a reality' {quoted on p. 35}). Ranjan forgets the last clause which says that appearance is reality, but what he wants to argue is made clear just one paragraph above this quote. If appearance were to be identical with reality, appearance would cease to be considered an appearance. This is then nub of his tirade against identity. Why then invite such a distinction first of all as a point of refutation against Churchland? Ranjan charges Searle: Searle does not know how to take appearance as only an appearance. Ranjan's position is now crystal clear. Churchland allegedly holds the following set:

(1) my experience of heat is a 'rough' feeling of neuron firings$_1$;
(2) my experience of heat is a vague feeling of neuron firings$_1$;
(3) my vague feeling of neuron firings$_1$, is the appearance of neuron firings$_1$;

(4) my vague feeling of neuron firings$_1$ is not the real neuron firings$_1$;
(5) but the real is identical with neuron firings$_2$ (Churchland's proof);
(6) since neuron firings$_2$ is as vague as neuron firings$_1$, appearance of neuron firings$_1$ is identical with the appearance of neuron firings$_2$.

What is the appearance of neuron firings$_2$ is the appearance of? Thus we are launched into an *infinite regressus.* Churchland's suggestion, coupled with the identity theory, leads to absurdity (p. 35).

What it proves is that while Churchland fails to grant ontological status to appearance, Searle does it in exactly the manner as our ancient Cārvaka materialists did. Hence there arises a case for comparative study of Searle and Cārvaka. But if this is the only basis for comparison, one must be told here that the sense in which it is understood by Searle is of comparable interest to Carvaka. This Ranjan cannot show by simply offering a formula of elements x, y and z and dubbing mind as a combination of them, but different from each one of them. In fact that Searle and Carvaka diverge in their thinking can easily be shown. What Searle hypothesizes is that water and H_2O are different 'aspectual shapes' of one and the same phenomenon, Carvakas tell us that mind is an emergent (epiphenomena?) or better put, self is the body (i.e. body is the substratum of self, and it follows therefore, consciousness is not something over and above the living body. In order to maintain ontological independence, this will not do. This is what Ranjan agrees to call as 'a brand of identity theory'. Where is the conceptual parallel, Ranjan claims here? On Searle's view, there is something irreducibly mental, consciousness is essentially subjective, and hence the ontology of the mental is essentially first-person ontology (*a fortiori,* materialists are wrong). Wherefore, the identity theory in Searle? I am not saying that there is no parallel, but it must be shown how a comparative study arises on the basis of the above. Ranjan has not yet made even a start.

Similar misgivings about Karl Popper's understanding of the puzzle about the relation between evolution and mind-body identity show that he quotes out of context and distorts their views. Popper does not claim to be an authority on cognitive science, but his remarks are directed against the emergent character of mind, and thus it adds support to identity theory. It is not poised to attack identity and therefore does not run counter to it. If Darwinism is causal-explanatory ('joint claim'), then it must also explain the reducibility of mind; but since it cannot receive any explanations, Darwinism's claim for causal explanatoriness may not be granted. Granting that the evolutionary explanation is adequate, then it only explains the emergent character, and therefore, it cannot explain in the way physical explanations do. It is more against its explanatory power than against psychological explanations. Mind cannot be explained as a product of evolution. Ranjan is misled by Popper's three Worlds. The causal efficacy Popper is speaking of is between World 2 and World 3 on

World 1 and he overlooks what Popper calls the 'joint claim' and Popper's conditional clause 'how little is said'. So his arguments on pp. 36–37 is a total distortion. Popper makes no assumption that mind is a product of evolution as assumed in

(b) in both of the arguments and hence it is wrong.

That brings me to his final word on unity, or the AP which goes as follows:

> For any Z, if there exists K such that Z is different from K, and K is different from Z, then Z is not ontologically self-existent is the concise statement of our 'exit' out our conceptual difficulties (p. 47).

On Ranjan's view, there is a deeper underlying unity which is ontological. This is what is given by the dream argument: Indian philosophical traditions recognize three stages of waking, sleep and deep sleep. Ranjan's argument assumes that there is a locus of awareness in dream (that will simply presuppose that our dreams are experiences; we can grant it though with great reservation for if dreams are experiences, the onus is to explain what kind of experience it is). Now dream experience ('awareness') and dream objects are identical (especially in deep sleep). (What empirical grounds are there for this assumption?) But still we cannot ascribe self-existence to dreamers and hence the awareness is different. This is the essence of his AP supported by 'principle of co-arising'. Immediately, Ranjan switches over to saying that dreams are not experiences because they cannot be positively characterized (What about REM, called so by rapid eye movements; but still the question whether dreams are experiences no doubt persists). But how does it follow that dream is existence itself from the above premises? What empirical or conceptual grounds are there, not only for the above, but also what justification is there for extending it to waking (then why make a distinction between waking and sleep; what is analyzed in sleep is better analyzable in the waking state). Norman Malcolm[2] has conceptually reduced dreams to absurdity. Daniel Dennett[3] has questioned the possibility of psychological explanations and concluded that dreams are the way we recall happenings back from our memory while we are awake, and *a fortiori* dreams are not experiences in yet another sense. What kinds of bearings this may have on Ranjan's remarks is anybody's guess. Ranjan argues: since I am identical with waking, body and mind are identical. What kind of argument is this? How the above premise entails the conclusion must have obviously escaped the notice of the editors. I think, much water has flown under cognitive science to believe what Schrodinger has said on identity is true today. From his premise about the singularity of consciousness (it is not experienced plurally), how does it follow that mind and body are identical? I think I have done much to expose the hollowness by reformulating the terms of his argument. I only hope

Ranjan himself will realize that Churchland's logic is not as quixotic as his, and the need to play Sancho Panza is not at all warranted by any of the above thinkers.

NOTES AND REFERENCES

1. Ranjan Umapathy, 'The Mind-Body Problem: A Comparative Study', *JICPR*, 1996, pp. 25–51.
2. Norman Malcolm, *Dreaming*, London, 1959.
3. Daniel Dennett, 'Are Dreams Experiences?' in *Philosophical Review*, 1976, pp. 151–71.

Department of Philosophy, University of Calicut A. KANTHAMANI

A Note on Sundara Rajan's Philosophy

It is over twenty years since Professor R. Sundara Rajan (RS) published his first major work, *Structure and Change in Philosophy* (1974) and has been writing till his passing away in June 1997. His major publications include: *New Studies in Marxism, Studies in Phenomenology, Hermeneutics and Deconstruction* (1991) and the trilogy, *Innovative Competence and Social Change*, (1986), *Towards a Critique of Cultural Reason* (1987), and *The Primacy of the Political* (1991), *Transformations of Transcendental Philosophy* (1994), and *Humanization of Transcendental Philosophy* (1997). Apart from these, atleast two of his works are almost ready.[1] He was very much philosopher's philosopher, eminent in his field, highly respected in academic circles. A close examination of his philosophy will show three strands in him: the phenomenological, the linguistic and the ecological. It is important to note that all these three are interconnected in almost all his writings. For example, in the triology, he has been insisting how a political community borns out of communication. Also, he has shown the inner connections and the interdependence between language and politics by discussing the political dimensions of language and the communication of discursive aspect of politics.[2] Similarly in *Studies in Phenomenology, Hermeneutics and Deconstruction*, he argued in support of phenomenological and hermeneutical problems. The problem of reference has been discussed at length mainly from the standpoint of Ricoeur and Derrida. Another significant contribution here is about his discussion on nature and Life-world and towards the end of the chapter he has examined the need for an ecological hermeneutics.[3] This has been again stressed by him in one of his excellent articles on 'Philosophy as Geo-Philia'.[4] One can easily say that RS's approach to philosophical problems has been always novel and analytic. What follows is a brief analysis of his approach which is reflected

in his paper, "Notes Towards a Phenomenology of Historiographies".[5] His attempt towards the phenomenology of historical writing and his conception regarding the three turns, namely the linguistic, the ecological and the feminist are worth analyzing. Moreover in the above paper, one can see the crux of his philosophy. In my short paper, I have attempted to throw some light on some of the points raised by him. I proceed as follows.

RS's main objective of the paper is to discuss the phenomenology of historical reason which he is able to achieve through Ricoeurean model (p. 187). Taking clue from Ricoeur, he argues how history could be either a series of actions or an account of record of such happenings. History, according to RS, deals with basic tendencies and dispositions of human subjects and the discursive practices. It is here RS argues that history has link with the ontological and hermeneutic level, on the basis of which he tries to construct the 'ontologically founded phenomenology of discursive practice', (p. 187) following Ricoeur. Thus he works on the direction to establish a phenomenology of historical reason. History in three perspectives are well approached by RS. He applies Holton's understanding of scientific theory in terms of three dimensions. The thematic dimension of Holton's understanding according to RS, is something new. What RS tries to establish here is to show that there are three such thematic principles, as power, idea and value, thus giving three kinds of historiography as (1) historiography as power, (2) as reason and (3) as vision. RS makes an attempt to show that in each thematic principle, there is a specific placement of the other two. For example, he says: 'Within the historiography of power, there is a specific placement of reason and vision. These other themata in turn seek to frame each other' (p. 192). I think, here RS is not maintaining any hierarchy among these and is trying to show that in each level, there is operation of other two. Perhaps, RS wants to maintain the view that though each is separate, that the role or part played by the other two are important and hence we cannot have a historiography of one dimension alone. Though RS's views on these three thematic principles are interesting, in reality, it is the historiography of power which is always a dominating factor. Besides this, the role of reason and vision can always be questioned and post-modernism has infact showed us the limitations of reason. It is true that the theoretical discourses of modernity from Descartes through the Enlightenment, championed reason as the source of progress in knowledge and society, as well as the privileged locus of truth and the foundation of systematic knowledge. In *Dialectic of Enlightenment*, Horkheimer and Adorno describe how reason can turn into its opposite. There are critics of modernists who argue that starting from the philosophical project of Descartes to the social theory of Comte, Marx, Weber and others, could easily be criticized for its search for a foundation of knowledge, and for its universalizing and totalizing claims.

In addition to this, there is also 'allegedly fallacious rationalism'. The interconnection which RS mentions, I think, may not be an acceptable and a necessary one. Such categorization will not help us in any way. In fact, it complicates the understanding process.

RS very nicely presents how the Heideggerean account of authentic historicity expands the conception of authentic agency. He further explains how we draw guidance from the past (p. 191) and provides an account of action as the transmission and realization of a tradition. The distinction RS makes with regard to both the authentic and inauthentic form of historical perception is important. He says: 'In authentic present, we have *responsibility*, while in the inauthentic mode, we have a *mastery*' (p. 191). Also by quoting Gadamer, RS explains the role of the past both in the present and in the future. Here one could see the influence of Heidegger and Gadamer on RS. But the danger of being dependent on the past has been emphasized by Habermas. Nietzche's classification of historiography into the monumental, the antiquarian and the critical, may help one here. One can see the parallels between the first two of these with the Heideggerean authentic historiography where there seems to be a dependence and reverence for the past. The authentic historiography of Heidegger simply allows historiography to be simply dependent of the past and there is no place of any critical. What happens is that we are overwhelmed by the happenings of the past and there is only reverence which does not allow any 'judging'.

It appears to me that RS has not sufficiently stressed the role of power, when he deals with the understanding of history, in the background of power, knowledge and good (p. 194). The role of all these has been taken into consideration to show the interconnection among them, but it has to be noted that among them, it is power which ultimately has been playing a dominant role in history. This point requires more of our attention because the role of ruling class and its power cannot be so easily neglected. The other classes not only live within the social arrangements made by the ruling classes' power, but also their very apprehension of that arrangement. The way they think about themselves and the social position are simply given to them by the dominant class. For example, Steven Lukes classifies power according to its three dimensions. In the first one, the idea of exercising power is to prevail over the contrary preferences of others; in the second, power serves as controlling the agenda, thus deciding what issues will ever get staged as matters to be contested or decided and the third dimension which not only includes the first two kinds, but also allows that power may operate to shape and modify desires and beliefs in a manner contrary to peoples' interests. The role of power has been stressed by Foucault also. For him, power has three characteristics. Power is productive; it is only exercised by individuals but never possessed by them; and it is involved in every social relation. Foucault makes it clear that the individual does not stand apart from

power or prior to it. The individual is constituted by power and hence, individual existence and identity are among power's effects. The individual never possesses power to act on his own. The individual, in the opinion of Foucault, exercises power at certain times and in certain places as a functionary of power's intentions, but not of his own. Thus Foucault concludes that 'a society without power relations can only be an abstraction' and that 'there are relations of power throughout every social field'.

Another significant contribution of RS in this paper is, about the threefold transformations, which he talks about, namely the linguistic, ecological and feminist turns. The linguistic turn has been, one of his major focus of attention in all his writings. For example, he very rightly argues that in linguistic turn, language is viewed not as a problem, but as a resource (p. 195). For him, it is a new way of understanding language itself (ibid). One could see the role of linguistic turn in RS's other works also. It is not clear whether RS supports the structuralists' position wherein it is held that it is not man who speaks language, but it is language who speaks man. Though nobody questions the performative role of language, it has to be considered only as a tool of understanding. The relation between nature and history comes under the ecological turn. In RS's writings these two turns have been playing a prominent role. But what is interesting is, now RS has moved to the third turn, namely, the feminist. Though very little has been said by RS about his views on the feminist turn in this paper, his view that it is a 'new way of thinking about all problems including those of men' (p. 198), has to be taken seriously. The paper ends with a positive note that supports the cause of unity of mankind. It is high time for us to work for such a new historiography, as suggested by RS.

Notes and References

1. R. Sundara Rajan has completed the work entitled, *Beyond the Crisis of European Sciences* (Vols 1 and 2). Vol. 1 has been completed by him when he was a Fellow of Indian Institute of Advanced Study, Shimla (1994–96). Vol. 2 has been completed when he was a Fellow of ICPR. (From 1996 to his untimely death). Both the volumes are yet to be published.
2. See his book, *Primacy of the Political*, Chapter IV.
3. See his book, *Studies in Phenomenology, Hermeneutics and Deconstruction*, Chapter III.
4. *Indian Philosophical Quarterly*, Vol. 22, No. 4.
5. *Journal of Indian Council of Philosophical Research*, Special Issue, June, 1996.

Ashok Vohra on Wittgenstein's Philosophy of Mind

In this discussion my aim is to evaluate Dr Ashok Vohra's 'Wittgenstein's Philosophy of Mind' published by Croom Helm, London and Sydney, 1986. Though a decade has passed for the publication, there is hardly any reaction in India on his book. This is an unfortunate condition of our academic philosophical world. Perhaps publication by a foreign publisher is not sufficient enough to attract the Indian philosophers. Vohra should have taken his birth in London or Sydney. He should have been non-Indian by birth. We are internationalist first and the nationalist never. What happens in this nation is not our concern. Though Vohra is perhaps the youngest Indian Wittgensteinian, he is mature enough to handle issues connected with Wittgenstein's philosophy of mind. I am not planning to produce a general review of his work, I would like to react on some specific issues. My critical approach towards his work does not at all mean that I do not regard his work as a remarkable achievement on the part of an Indian Wittgensteinian.

There are several ways in which a philosopher can be handled. His views, his idiom of expression etc., may not be familiar to the people. So the unfamiliar views and unfamiliar idiom have to be explained in terms of the familiar views and familiar idiom. The assumption is that Wittgenstein's views, his idiom, is not known to the people. Vohra's job is to make people acquainted with Wittgenstein's views and idiom. To do this one way could be that of a purist, one who does not take help of other philosophers, whether well-known or comparatively not so well-known. The other way could be that of an impurist who takes the help of other philosophers in order to explain Wittgenstein's views. I prefer the way of the purist. The way of the purist does not have the danger of looking at Wittgenstein in the image of some other philosopher. Vohra has preferred the way of the impurist. He has brought to our notice so many philosophers of our time; the prominent among them are Ryle, Strawson, Ayer and Shoemaker. Obviously, he took for granted that their views are well-known to the people. Unknown is to be explained in terms of the known. People are to be familiarized with Wittgenstein's views with the help of the philosophical views with which they are already familiar.

Vohra, while writing about Wittgenstein in terms of Ryle's views, found himself in a paradoxical situation. He discovered that he has now an added burden. His project required that he should explain the position of Ryle first. Vohra's book has only four chapters, and the first chapter, which is supposed to be an introductory chapter in any book, is exclusively devoted to Ryle. Wittgenstein's name does not occur anywhere in the whole chapter. So also not a single work of Wittgenstein has its reference in the 'Notes' on the chapter. A purist would really be shocked. How could Vohra manage to write on Wittgenstein's philosophy of mind in

which the introductory chapter is devoted to Ryle's philosophy of mind? It is not only the first chapter, Ryle is the moving spirit of the whole book, followed by the smaller spirits like Strawson, Ayer, Shoemaker and others. Wittgenstein's name has receded to the background. Even in the third and the fourth chapters a very few citations from Wittgenstein have occurred which will be discussed later. The proper title for Vohra's book should have been 'Philosophy of Mind: Perspectives of Ryle, Strawson, Shoemaker and Wittgenstein'. Wittgenstein's perspective cannot be avoided because the second chapter on Private Language has so many references to Wittgenstein. But even the second chapter has ended up with Ryle's jargon. 'Those who allocate the concept 'private' to the concept 'language'', Vohra maintains, have 'made a category-mistake.'[1] Category-mistake is Ryle's jargon. Wittgenstein would certainly reject this jargon. Consider the sentence 'vowel *e* is yellow'. To Vohra this sentence may raise a category mistake. But for Wittgenstein it exhibits no kind of mistake. Referring to this sentence Wittgenstein remarks, 'I could not express what I want to say in any other way than by means of the idea 'yellow'.'[2] Solipsism is certainly not devoid of all sense. Otherwise how could people write so many books and articles on it? And for a solipsist the language is necessarily private, he commits no category mistake about it. Of course he may be committing a philosophical blunder. But a philosophical blunder is not a category-mistake. Not only early, even later Wittgenstein committed this blunder.

By seeing the title of Vohra's book, one would start expecting that the book would be crammed with references to Wittgenstein; it would be overflowing with Wittgensteinian landscapes. But Wittgensteinian landscapes are rare. Wittgenstein has been presented as an approver, now of Ryle's view, now of Strawson's view, now of Bernard Williams' view and so on. It seems as if Wittgenstein has no identity of his own; his identity is merged into the identity of so many British philosophers. Consider the sample. Vohra writes 'As "I" does not have a stable, but a shifting referent, it can, therefore, be called an "index-word". Ryle likens it in this respect to a word such as "now" which also can be applied to any point whatever in the time series. Wittgenstein expresses the same fact by saying that "I" is not the name of a person, nor "here" of a place, and "this" is not a name (*PI* 410).'[3] Wittgenstein does not reject 'I' as the name of a person because 'it does not have a stable reference or that its reference goes on shifting' but that it is not a referring expression at all. So Wittgenstein is not an approver of Ryle's view. Vohra is wrong in giving the testimony of Wittgenstein in support of Ryle's view. Though Wittgenstein denies that 'I' and 'this' are names, he does not deny their connection with names. 'Names are explained by means of them.'[4] How can they explain names if they are just index-words?

Not only Wittgenstein, Vohra's other favourite philosopher Strawson shares the same fate as that of Wittgenstein. He too has been presented

as an approver of Ryle's view. Following Ryle, Vohra considers mind as a 'blanket concept' like the 'British Constitution'.[5] Ryle's reflections on these two blanket concepts have led Vohra to infer Strawson's concept of a person which is accepted by him as his own concept, as his own view. Questions concerning the relation between mind and body, according to Vohra, are improper questions. 'They are improper in much the same way as is the question, "What transactions go on between the House of Commons and the British Constitution?" In our view, therefore, the concept of a person is a primary concept. By saying that the concept of a person is a primary concept we mean: the concept of a person is not to be analyzed as that of an animated body or of an embodied anima. This is not to say that the concept of a pure individual consciousness might not have a logically secondary existence.'[6] Vohra continues quoting Strawson. It does not occur to Vohra that Ryle is playing one game and Strawson another. Ryle's objective is not to establish that the concept of a person is a logically primary concept. His objective is simply to drive the Cartesian ghost away from the body. Strawson does not drive the Cartesian ghost away from the body; he simply degrades its status; he lowers its prestige. Strawson does not find any incoherence in the concept of an 'embodied anima'. No mistake is committed in bringing the two diverse concepts together. It makes quite good sense to talk about an embodied anima. But such a talk presupposes the talk about a person. Talking about a 'disembodied person' or an 'embodied person' is not a nonsensical talk, but a logically secondary talk.

At times Vohra seems not even to remember what he did and what he did not do in a certain context. For instance, he begins his second chapter by writing, 'We established in the previous chapter that we need not posit any private entities like sensations, sense impressions or sense-data to account for our observation.'[7] But in the 'previous chapter', i.e., the chapter first, he rejected neither sensations nor sense impressions. So also he has not attacked their *privacy*. He has not shown that they are posited as *public* objects. He has only attacked sense-data, and his attack suffers from the limitations of Ryle. Consider first Vohra's views about sensation and sense impressions. He has simply summarised the views of Ryle on them. Vohra quotes Ryle with approval 'Having at least one sensation is part of the force of "perceiving", "overhearing", "savouring" and the rest.'[8] And concerning sensations Vohra goes on saying such things as 'Sensations cannot exist unfelt or unhad.'[9] 'Nobody, however competent or well placed he may be, can feel my sensations.'[10] 'Nobody except myself can feel my sensations.'[11] These are Vohra's own views about sensations established in the first chapter of his book. They are not the views of others quoted for comments and analysis. Of course they are Ryle's views which Vohra has owned without any hesitation. Not only Vohra posits sensations, he is their exclusive owner; they are his private property. Vohra's position is that sensations must be posited, and posited

in the private sense, in order to have perception or observation of horse-races and the races of all other kinds.

The position with respect to sense impressions is no better than the position with respect to sensations. Vohra makes such remarks about sense-impressions as 'About both sensations and sense impressions we say that we have them.'[12] 'We cannot be mistaken about both.'[13] 'As with sensations, I alone can have sense impressions.'[14] 'Like sensations, sense impressions cannot exist unhad.'[15] Thus like sensations Vohra secures privacy also for sense impressions. What Vohra denies is simply having a glimpse of a sensation or of a sense impression. They are not the sorts of things of which one can have glimpses. Of course the level of sense impressions in the account for observation is different from the level of sensations. Sensations belong to a more basic level than sense impressions. A sensation is a constituent of observation, it makes observation possible, therefore, it cannot occur as an *object* of observation. As Vohra remarks 'If sensations were the proper objects of observation, then observing them must entail having at least one glimpse of those sensations. . . But this is absurd.'[16] All this has led Vohra to produce the formula 'Observation=Having a sensation + Paying heed to it.'[17] Obviously Vohra does not mean that in order to observe something one has to attend to one's sensations. If one attends to one's sensations, if one pays heed to them, then one may not pay any heed to the object that lies outside one's sensations, the object that awaits observation. Vohra's formula does not lead us to the 'object of observation', it misleads us; it directs us to attend to our sensations. You cannot ask Vohra what he means by 'paying heed to', because according to him the question 'How does one pay heed?' is a 'spurious question' for which 'no answer can be given'.[18]

The unfortunate situation is that Ryle would himself reject Vohra's formula. Though at the initial stage he accepted that sensations are the basic ingredients of observation, later he came to accept that sensations, instead of helping us impair our observation. 'Sensations do not help, they hinder perception. I see much better when I am not being dazzled than when I am. Sensations, in this sense, are not usually present when perception occurs; and when they are present they tend to impair perception. They are not *sine qua non* of perception."[19] Obviously Ryle means that the sensations which are *sine qua nons* of perception are only technological inventions. Vohra sticks to Ryle's earlier view otherwise he would have failed to produce the formula that he produced.

Consider now Vohra's attack on sense-data. As has already been pointed out that Vohra's attack on sense-data suffers from Rylean limitations. Vohra has tried to show that the direct objects of perception are *material.* It is wrong to hold that sense-data are the direct objects of perception. Though we do not perceive material objects in dreams, hallucinations and illusions etc., this does not mean that we perceive some other kinds of objects in them, called sense-data. As Vohra points

out 'when we suffer from an illusion there is no object at all, physical or non-physical, which we are perceiving in any possible sense of "perceiving". We are simply under the false belief that ordinary perception is taking place. The difference between ordinary perception and sensory illusion is that in the case of ordinary perception the beliefs are true, whereas in the case of sensory illusions the beliefs are false.'[20] Sense-data philosophers may be wrong, but Vohra is certainly not right. For Vohra has assigned truth-value to a belief in terms of 'normal perception' and 'illusory perception'. A true belief coincides with a normal perception, and a false belief with illusory perception. It is impossible to have a false belief when one's perception is normal. Similarly, it is impossible to have a true belief when one's perception is illusory. How do we know that a perception is normal? It will be known if one knows that the belief connected with it is true. And how does one know whether a belief connected with a perception is true? It will be known if one knows that the perception connected with the belief is normal. So assigning truth-value to beliefs on the ground of 'normal perception' and 'illusory perception' would suffer from vicious circularity. Normalcy of a perception would depend on the truth of the belief connected with it, and the truth of the belief on its own part would depend on the normalcy of the perception connected with it. In ordinary discourse 'normal perception' does not exclude the possibility of a false belief, hence the ordinary discourse avoids the vicious circularity in question.

The other argument of Vohra against sense-data is stronger than the argument mentioned above. Vohra argues 'we can only analyze perceptions in terms of theoretical entities—e.g. "?sense-data"—after we have already perceived those objects. That is, the theoretical analysis of the perception of physical objects in terms of "sense-data" is dependent on the prior perception of those (physical) objects.'[21] Unless physical objects are perceived, and are believed to exist, one cannot proceed to produce the sense-datum analysis of them. Sense-data philosophers would have hardly any disagreement with Vohra on this issue. Vohra does not seem to be aware of the discussion of sense-data after the popularity of Paul's paper on sense-data.[22] The whole discussion of sense-data took a new turn because of Paul. Sense-data philosophers stopped denying the existence of material objects, so also they stopped refusing to grant proper status to material object language. Consider some remarks from the introduction of a later edition of Price's *Perception.* In order to understand the terminology of sense-data, according to Price, 'the material object language of commonsense must be understood already.'[23] This leads Price to the question 'How can we have learnt to understand it, unless some material object words have ostensive definition?'[24] So sense-datum language, according to later Price, presupposes the existence of material object language. In a similar tone Ayer accepts, 'It is, in fact, only by the use of expressions which refer to the perception of physical objects

that we have given any meaning to talking of sense-data at all. . . If one has to describe the use of an unfamiliar terminology, the description, in order to be informative, must be given in terms of what is already understood.'[25] We are familiar with the terminology of material objects and unfamiliar with the terminology of sense-data. The unfamiliar terminology is explained in terms of the familiar one. Paul brought a change in the views of Price and Ayer. According to Paul, 'Sense-data are verbal novelty rather than factual discovery'. 'To talk about sense-data' is just another way of 'talking about the way objects look.'[26] Paul finds the sense-datum terminology merely as an alternative to physical object terminology. 'We can say a thing this way and we can say it that way.'[27] Then why should one opt for the sense-datum terminology? Because, this terminology avoids reference to hidden substances of the physical object terminology. Vohra has not cared to see the post-Rylean literature on sense-data.

Vohra is best in his second chapter on 'Privacy and Private language.' He is a serious Wittgensteinian in this chapter. However, even in this chapter one can find glimpses of Ryle. These glimpses have damaged his vision about the Wittgensteinian solution to philosophical problems. Consider one such instance. Vohra writes 'The case of pains is not like physical objects, but is like colours. To assimilate pains to physical objects is to make a category mistake.'[28] For Vohra there is no category mistake if pains are assimilated to colours, but there is a category mistake if they are assimilated to chairs. Pains, colours and chairs are on their faces diverse concepts. If pains can be permitted to go with colours, how can they be stopped from going with chairs? They can be stopped because Vohra gives quite diverse interpretations to colours and chairs. He argues 'When we talk of the identity of physical objects we talk of two or more things. But, in the case of the identity of colours, when we say, "This colour here is exactly like the colour over there", or "This is the same colour here as over there", whichever of them we say, there is but one colour—say, red—and it would be a mistake to say "There cannot be only one colour; for there is a colour here and also that colour there."'[29] Vohra's argument demonstrates half-digested Platonism. A chair in *this* place is numerically different from a chair in *that* place, but a red-patch in this place is *not* numerically different from a red-patch in that place. Why? One can very well argue that one and the same chair is occurring in two different places. A chair here is numerically the same as a chair there, just as a colour-patch here is numerically the same as a colour-patch there. If it is impossible for a chair here to be numerically the same as a chair there, then it is also impossible for a colour here to be numerically the same as a colour there. Both of them, colour and chair, stand and fall together.

Pains do not behave differently from chairs and colours. This is the

position of Wittgenstein, and not the one for which Vohra has argued. Consider the chair placed in this room. It would remain numerically the same chair inspite of the fact that numerically different persons, say 50 persons, have perceived this chair. Colour behaves in the same fashion as a chair. A colour-patch occurring in this locality would remain numerically the same colour-patch even if numerically different persons perceive it. The same is true about pains. Suppose a toothache occurs in a certain locality, the locality could be the mouth of a person *x*. Is it ruled out that the persons *y* and *z* locate their own toothaches in the same locality? When asked where the ache occurs *y* brings his finger to the mouth of x and identifies his toothache in the tooth of *x*. The same thing happens with *z*, he too brings his finger to the tooth of *x*. In such a situation we would be led to say that numerically the same toothache is felt by three different persons. This situation is similar to the situation in which the same colour patch or the same chair is seen by three different persons. Like the numerical identity of a colour-patch or the numerical identity of a chair, the numerical identity of a toothache depends on location. If the locations of toothache differ, then numerical diversity would be the consequence. Suppose the toothache occurs in three different localities, one locality is the mouth of *x*, the other the mouth of *y* and the third, the mouth of *z*. Then there are numerically three different toothaches. This is like chairs and colours, three different chairs occurring in three different rooms and three different colours attached to three different material objects. This is how Wittgenstein distinguishes one toothache from the other. As he remarks, 'How are toothaches to be distinguished from one another? By intensity and similar characteristics, and by location.'[30] Wittgenstein certainly does not mean that one and the same toothache occurs in all possible mouths. A toothache occurring in one mouth has a different location from the toothache occurring in another mouth. But the location criterion has its limits.

Sometimes, though not always, the location criterion breaks down. Though three different people may trace their toothache in the same locality, the mouth of *x*, the toothache of *y* may be very intense, it may be unbearable. In such a situation we would be led to say that *y*'s toothache is numerically different from the toothaches of *x* and *z*. The situation is similar to a colour patch. A normal person sees a red patch which a colour-blind person sees as black patch. In such a situation we would be led to say that they see different colours and not numerically the same colour. Though the location of colour-path is the same; in this same location they see numerically different colour-patches. Similarly, though the location of toothache is the same, *y*'s toothache is numerically different from the toothache of *x*. For his toothache *y* requires immediate attention of the doctor, but *x* is in no hurry.

Vohra has certainly been misled by Wittgenstein's remark. Wittgenstein says 'in so far as it makes *sense* to say that my pain is the same as his, it is

also possible for us both to have the same pain. (*PI* 253).'[31] Vohra is led to think that the toothache that occurs in one locality (the locality called 'my mouth') has to be numerically the same as the toothache that occurs in another locality (the locality called 'your mouth'). Toothache has to remain the same in spite of its occurring in different localities. This imagination has led him to compare pains with colours, and further led him to introduce different criteria for the identity of colours from the identity of chairs. But if I trace my toothache in the same locality as you trace your toothache (it may be 'my mouth' or 'your mouth'), then it makes quite good sense to say that my toothache is the same as your toothache, that we are not having different toothaches, we have numerically the same toothache. Wittgenstein does not at all intend to abolish the numerical differences between different toothaches. Toothaches differ from one another as colours differ from one another and as chairs differ from one another. No category mistake is committed by assimilating pains to colours and chairs.

Suppose the toothaches are distinguished from one another, not in terms of their locations and intensity etc., but in terms of the persons who *own* them. Wittgenstein rejects such a suggestion by saying 'But if it is objected that the distinction is simply that in one case *I* have it, in the other *he*; thus the owner is the defining mark of the toothache itself; but then what does the proposition 'I have toothache' (or someone else does) assert? Nothing at all.'[32] A definitional truth is not a true/false assertion.

Wittgenstein has simply contemplated a possible philosophical alternative to his own view and rejected it. But this alternative has seriously been taken up by Strawson. Strawson maintains 'states, or experiences, one might say, *owe* their identity as particulars to the identity of the persons whose states or experiences they are.'[33] Not only this, there is also a requirement for 'an independent identification of the sufferer of the experience.'[34] This means that in order to identify a toothache one must identify the person who owns it. But any attempt at identifying the owner presupposes a doubt about his identity. It is quite interesting to note Wittgensteinian's reaction to Strawson's alternative. Wittgenstein reacts, 'There is no question of recognizing a person when I say I have toothache. To ask "Are you sure that it's you who have pains?" would be nonsensical. . . It is as impossible that in making the statement "I have toothache" I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me.'[35] Since I have *failed* in identifying myself I have also failed in becoming the *owner* of the toothache in question. I do not become the owner of a toothache simply because I *felt* it, anymore than I become the owner of a colourpatch simply because I *saw* it or the owner of a chair simply because I had a *visual glimpse* of it. Feelings and visual glimpses establish a contingent relation between me and toothache, or a colour, or a chair.

This very toothache, this very colour, this very chair may have contingent relation with others. They may desert me.

Vohra's third chapter creates more difficulties of comprehension than his earlier two chapters, and these difficulties continue till the end of the fourth chapter. Consider the difficulties. Referring to the 'substance' and the 'bundle' theories Vohra says 'Both these theories are based on the assumption that all mental happenings are ghostly episodes, which are the work of a ghost, namely the mind, whose home is in the body.'[36] By the 'substance theory' Vohra means the theory which is associated with the name of Descartes, against which Ryle invented his famous abuse 'Dogma of the ghost in the machine'. And this abuse has now become the legacy of our age. By the 'bundle theory' Vohra certainly means the theory that is associated with Hume's name. Vohra has given a quite unusual interpretation to Hume's view, an interpretation which may surprise some and shock others. The usual interpretation is that Hume did not allow even the 'formal structure' of the ghost to occur. He refused even formal unity to experiences. Hume left this job for Kant, therefore Hume's view is described as the bundle-view of experiences. Where is the question of accepting the existence of a body for Hume? How can it be shown that Hume believed in the existence of a 'mind whose home is in the body'. And if Vohra succeeds in showing it then he has proved that Hume was Descartes in a new guise. Descartes has taken birth with the thought-form of Hume. It is both strange and shocking that the 'bundle theory' should be considered as a new name for the 'substance theory'.

Though Vohra has written on Wittgenstein's Philosophy of Mind, a very few remarks of Wittgenstein have been cited in the third and the fourth chapters. These chapters deal with such interesting issues as self-knowledge, personal identity and other persons. These chapters are crammed with references to other philosophers and their works. By isolating the remarks of Wittgenstein from these chapters we can see whether these remarks contribute in any significant way to the issues in question. In the third chapter there are only five references to Wittgenstein's work, and all are from *Philosophical Investigations.* One of them has already been discussed, the remark that '"I" is not the name of person, nor "here" of a place, and "this" is not a name. (*PI,* 410).'[37] This remark has been quoted in support of Ryle. All the other four remarks are connected with the development of what Vohra describes as the 'formal features of our lives'. And in doing this he thinks that he is supporting Wittgenstein and 'disagrees with Ryle".[38] Exhibiting his disagreement with Ryle and agreement with Wittgenstein Vohra remarks 'I agree with Wittgenstein and hold that such psychological states as being in pain, being sad, being happy are the formal features of our lives, and the first person expressions embodying them such as 'I am in pain', 'I am sad', 'I am happy' do not need any reasoning. They are simply the ways in which I act (cf. *PI,* 217). 'It is natural for us to say a sentence in

such and such surroundings, and unnatural for us to say it in isolation' (*PI*, 595). The whole idea of proof, or justification or reasoning about formal features of our life is confused one."[39] By 'formal features' Vohra simply means 'psychological features'. Of course, 'formal' is distinguished from 'contingent' or 'material'. So psychological features according to Vohra are essential or necessary features of our lives. Our life would be no life if we lack psychological features. What about our bodily features? They are not formal features, they are contingently related to us. They do not exhibit the essence of our lives. Essence of our life is exhibited by the psychological features. So Vohra has quite intelligently introduced the Cartesian ghost into Wittgenstein's philosophy. He forgot not only his Wittgenstein but also his Ryle. May be his opposition to Ryle is shown by converting Wittgenstein into a believer of the ghost in the machine, and then siding with Wittgenstein. But the difficulty does not end here. There is a sense in which one can see Wittgenstein also in the image of Hume. If one's essence consists in having one or the other psychological state, being happy or unhappy, being in pleasure or pain etc., then one is a Humean. Vohra has interpreted Wittgenstein in such a fashion that at one time he appears as a Cartesian and at the other as a Humean. Vohra's picture of Wittgenstein's view is not very unlike Jastrow's duck-rabbit picture. One moment it appears as a duck's head, and at other moment as a rabbit's head.[40]

There simply exists no puzzle about Wittgenstein's view. Wittgenstein was neither a Cartesian nor a Humean. He attempted to de-psychologize the first person psychological statements. The analysis of 'I have a pain' in terms of 'There is a pain' is an attempt in this direction.[41] So also his attempt to develop the notion of a pain-patch on the pattern of a colour-patch is a further step in this direction.[42] What Vohra describes as the 'formal features of our life', Wittgenstein wished to see them as the non-formal features of the world. A toothache for Wittgenstein has the same status as a colour-patch or a chair.

In order to show that the demand for evidence in the case of first person psychological statements is illegitimate Vohra has quoted *PI*, 381. He writes, 'To the question "How do I know that this colour is red?" the only answer available is: "I have learnt English".'[43] He has extended this reasoning even to such statements as 'This is a tree' and 'This is an inkpot'. But one has to be careful in extending any reasoning, including the one given by Wittgenstein. Suppose someone says 'I know this is arsenic', and the question is asked 'How do you know this is arsenic?' In such a situation would you be happy with the response: 'I have learnt English'. Take another example. A witness in the court announces 'I know he is a murderer'. The Judge asks: 'How do you know he is a murderer?' Would the Judge be satisfied with the response of the witness: 'I have learnt English.' Philosophers too are like judges; they do not pass judgements without evidence. When there is no evidence for the case, the case is rejected. First person psychological statements do not have evidence, therefore, they cannot be knowledge-claims of any kind. Then how can they function as the foundational claims in a theory of knowledge? In the structure of a theory of knowledge some claims may be more primitive than the others, but all of them have to be the knowledge-claims of some sort or the other. And a knowledge-claim is rejected not only because it lacks evidence, but also because it is not a socially accepted claim. As Wittgenstein remarks, 'I look at an object and say "That is a tree" or "I know that that's a tree" . . . But if all the others contradicted me . . . what *good* would it do to me to stick to my "I know"?'[44]

Consider now Vohra's references to Wittgenstein in the fourth, i.e., the last chapter of his work. This chapter has the title 'Knowledge of other Persons' which simply means the 'knowledge of other minds'. Referring to what Vohra proposes to do in this chapter he remarks 'I propose in this chapter to show that there is no epistemic problem about the knowledge of other minds, and their identity either.'[45] In support of his view Vohra has quoted Wittgenstein's remarks on the 'human soul' which simply do not support his view. Consider the remarks quoted by Vohra: 'My attitude towards him is an attitude towards a soul, I am not of the opinion that he has a soul', and 'The human body is the best picture of the human soul'.[46] Wittgenstein's remarks can help Vohra only if Wittgenstein's concept of the soul is equated with the 'concept of a subject of experiences.' In this sense if one is acquainted with a soul one is acquainted with a subject of experiences. But this does not help. For Wittgenstein clearly makes a distinction between 'attitude towards a soul' and 'attitude towards someone who has a soul.' Wittgenstein's context for the remarks on soul has been provided by religion, not by the theory of knowledge. Sandwiched between the two remarks on the soul quoted by Vohra are Wittgenstein's remarks on religion. He said such things as 'Religion teaches that the soul can exist when the body has disintegrated. Now do I understand this teaching?—Of course I understand it.'[47] Thus, there is every reason to think that when Wittgenstein refers to the soul while saying 'My attitude towards him is an attitude towards a soul', he has not reduced the soul to a subject of experiences. He does not mean to say 'My attitude towards him is an attitude towards a subject of experiences.' Before using Wittgenstein in support of his analysis Vohra should have shown that Wittgenstein's concept of the soul in this context means the subject of experiences. If Wittgenstein's concept of the soul, in this context does not mean the concept of a subject of experiences, then Vohra is not entitled to say such things as 'We do not attribute pains to stones, tables and chairs, because our attitude towards a stone and towards a being having human body is different.'[48] The attitude to which Wittgenstein refers is the attitude of a religious person. His problem is not the reduction of human beings to mere subjects of experiences, but making them free from this bondage.

There is no doubt that Wittgenstein does consider human beings as the subjects of experiences. And it is also possible that they could be such subjects because they *are* souls, or *have* souls. But their being souls is independent of the fact that they are subjects of experiences. Vohra is right when he refers to Wittgenstein saying: 'We only say of a human being and what is like one that it thinks (*PI*, 360).'[49] Once it is established that someone is a human being it is also established that he thinks. The epistemic problem of other minds becomes easy to solve. But Wittgenstein himself does not allow the problem of other minds to be solved so easily. Consider his remark: 'We don't say of a table and a chair that they think; neither do we say this of a plant, a fish, and hardly of a dog; only of human beings. And not even of all human beings.'[50] If all human beings do not think, then the problem of other minds remains where it was. Even if it is established that someone is a human being, it has yet to be established that he thinks. So Wittgenstein has provided no solution to the problem of other minds.

The argument from analogy for the existence of other minds has been as forcefully attacked by the philosophers of our time as it was supported at one time. Vohra has also contributed to this attack. It is in connection with his attack that he has referred to Wittgenstein. He explains, 'If I generalize from one case only, namely, my own, then it could only be called an irresponsible generalization', and quotes Wittgenstein saying 'How can I generalize the *one* case so irresponsibly? (*PI*, 293).'[51] But those who have attacked the argument from analogy, have never succeeded in producing a better alternate argument. They bring out this very argument, but in a round-about fashion by playing with words. Vohra is no exception. Vohra maintains, 'by virtue of the training we have received we know that the other walking talking figures which have bodies similar to mine, have minds.'[52] Vohra's steps during the training period can be explicated:

	Self-Knowledge	*Knowledge of Others*
1.	I am a walking-talking figure.	Smith is a walking-talking figure like me.
2.	I have a body which makes walking and talking possible.	Smith has a body like me.
3.	I have a mind that thinks.	Smith has a mind like me.

My training in knowing about myself and others has not been very unlike the training given to a 'chicken-sexer' to whom Vohra refers for explanation.[53] A chicken-sexer knows whether a chick is male or female. Similarly I know, because of my training, that I am different from others. But perhaps the chick knows, without any training, that it is different from the other chick which is a female. Could the case of a human child be different? Have not the philosophers unnecessarily involved themselves in weaving their theories about other minds. Perhaps their worries and

anxieties would have been avoided if they had contacted the human-chicks.

It is far more difficult to understand Vohra's views about personal identity than to understand his views about other minds. While introducing the issue of personal identity in the third chapter he gives the impression to his readers that he is rejecting both the theories of personal identity, the theory that advocates psychological criteria of personal identity and the theory that advocates the use of bodily criteria for such an identity. As he says about his project: 'In the main I have rejected the theories advocating the memory and bodily criteria of personal identity respectively.'[54] One may be little disturbed with this proposal. We are acquainted only with these two criteria. If Vohra is rejecting both the criteria, then he might be planning to introduce a third criterion, something brand new. But Vohra's proposal is not to increase the stock of criteria for personal identity. His proposal is simply to withdraw both these criteria. According to him 'we know about ourselves and about our identity without any theories.'[55] 'Without any theories' means 'without any criteria'. Theories introduce criteria. For knowing ourselves and our identity no criteria are needed.

It is quite shocking that Vohra does not keep his words. He demolishes all our expectations roused by him. Instead of rejecting both the criteria, as he proposed to do, he rejects the criterion of psychological identity and favours the criterion of bodily identity. Consider the last two paragraphs of the last chapter of his book. His conclusive decision is 'bodily identity can be regarded as a criterion of personal identity.'[56] In support of this view he says 'we pass judgements of identity on other persons on the basis of similarity of bodily appearance and other behavioural criteria.'[57] His last sentence of the chapter is 'The technique of knowing the identity of the other minds includes many observable facts about others, but the primary among them is the bodily appearance.'[58] Is there any doubt that Vohra has supported the theory of other minds that involves bodily identity as the criterion of personal identity? Then what right has he to criticize those who depend on some theory or the other for knowing the mind of others?

In the last I would like to say that my critical approach to Vohra's book should not lead one to think that his book lacks all virtues. It is full of virtues. The most important virtue for a philosopher is clear writing. Very few philosophers have this virtue, and Vohra is one among them. In the preface to his book he writes, 'In writing this book I have avoided confusing profundity with obscurity. I have employed a clear and simple conceptual apparatus, to present my point of view.'[59] Vohra retains this spirit throughout the book. If he had not been clear it would have been impossible to criticize him. Plato and Berkeley are rejected everyday, and that is possible because what they wrote, they wrote clearly. The other important virtue which the book possesses is awareness about the

contemporary thinking on a given philosophical problem. Vohra has clearly demonstrated that he is aware of the views of nearly all the important philosophers of our age who have written on the philosophy of mind. His scholarship cannot be doubted. So also his argumentative ability cannot be doubted. The book is full of arguments. No attempt has been made by Vohra to make his book thick by adding superfluous material. He has not started with the programme of writing a thick book. His programme was to write a good book; not necessarily a thick book. I feel that he has succeeded in his programme. He has made us aware about the difficulties connected with the issues of knowing one's own mind and the mind of others. Thus he has succeeded in performing the job of a philosopher.

Notes and References

1. Ashok Vohra, *Wittgenstein's Philosophy of Mind*, Croom Helm, London and Sydney, 1986, p. 66.
2. *Investigations*, II, xi.
3. Vohra, *Wittgenstein's Philosophy of Mind*, p. 81.
4. *Investigations*, 410.
5. See Vohra, *Wittgenstein's Philosophy of Mind*, pp. 76–77.
6. Ibid., p. 77.
7. Ibid., p. 34.
8. Ibid., p. 12.
9. Ibid., p. 13.
10. Ibid., p. 14.
11. Ibid.
12. Ibid.
13. Ibid.
14. Ibid., p. 15.
15. Ibid.
16. Ibid., p. 17.
17. Ibid., p. 12.
18. Ibid., p. 32.
19. Quoted by Vohra (p. 18) from Ryle's 'Sensations', *Contemporary British Philosophy*, Theory series, ed. H.D. Lewis, London, 1956.
20. Vohra, *Wittgenstein's Philosophy of Mind*, p. 22.
21. Ibid., p. 28.
22. G.A. Paul, 'Is There a Problem about Sense-data', *Proceedings of the Aristotelian Society*, 1936–37, reprinted in *Logic and Language*, Blackwell, 1951.
23. H.H. Price, *Perception*, edition of 1961, London, p. VIII.
24. Ibid.
25. A.J. Ayer, *The Problem of Knowledge*, p. 111.
26. See J.O. Urmson, *Philosophical Analysis*, Oxford, 1956, p. 185.
27. Ibid., p. 186.
28. Vohra, *Wittgenstein's Philosophy of Mind*, p. 50.
29. Ibid.
30. *Philosophical Remarks*, 61.
31. Vohra quotes *PI*, 253, in support of his view. See p. 50.
32. *Philosophical Remarks*, p. 61.

33. *Individuals*, Metheun, London, 1959, p. 97.
34. Ibid., p. 41.
35. *Blue and Brown Books*, p. 67.
36. Vohra, *Wittgenstein's Philosophy of Mind*, p. 75.
37. Ibid., p. 81.
38. Ibid., p. 80.
39. Ibid.
40. See *Investigations*, p. 394.
41. See *Individuals*, p. 95.
42. See *Investigations*, p. 312.
43. See Vohra, p. 79.
44. *On Certainty*, p. 503.
45. See Vohra, p. 95.
46. *Investigations*, Part II, p. 178. See Vohra, p. 106.
47. Ibid.
48. See Vohra, p. 106.
49. Quoted by Vohra, p. 103. Vohra omits the completion of the paragraph. Wittgenstein continues saying 'We also say it of dolls and no doubt of spirits too.' So thinking has diverse applications.
50. *Philosophy of Psychology*, Vol. II, p. 192.
51. Vohra, p. 99.
52. Ibid., p. 106.
53. Ibid.
54. Ibid., p. 68.
55. Ibid.
56. Ibid., p. 109.
57. Ibid.
58. Ibid.
59. Ibid., preface.

Hyderabad Suresh Chandra

Why Flog a Dead Horse !
A Response to Suresh Chandra's 'Ashok Vohra on *Wittgenstein's Philosophy of Mind*'

Suresh Chandra could not be more right when in the opening paragraph of his evaluation of my book *Wittgenstein's Philosophy of Mind*, he states: 'There is hardly any reaction in India on his book.' All these years I had a feeling that the book never existed, at least for the Indian colleagues in the profession. Though the book was reviewed in a few philosophy journals abroad, none of the philosophy journals published from India reviewed it, even though it was sent to them by the publishers. Some well-known Wittgenstein scholars discussed it with me, and some, including Professor G.H. von Wright wrote detailed notes to me about it when the book was sent to them by the publishers. Professor P.F. Strawson discussed it with me when he was in India, but no Indian scholar except Professor

Daya Krishna ever discussed it with me. In a country where oral tradition is still in vogue, and in which some people still believe that age rather than work and achievements should be the criterion for judging a person's worth, nothing more could be expected. So, when Professor Suresh Chandra's review of the book was brought to my notice I was rather thrilled and my belief that all is not lost for those who have a firm faith in the written word rather than in the oral tradition was reconfirmed. How true it is that the Indian academia in general, and Indian philosophers in particular, suffer from the malady of not taking appropriate cognizance of their compatriots!

At the outset I must say that I find it extremely difficult to react to the criticism levelled against me by Suresh Chandra. First, because though the book was published more than a decade ago, the work on it had started about a dozen years before that. The book, therefore, has to be evaluated in its own time frame. The thesis propounded or explained in it has to be seen and examined in its context, and the style of the book is to be evaluated keeping in view the paradigm available at the time of its writing. Looked at from the vantage point of today we may find it jejune and juvenile as it does not conform to the idiom currently in vogue in Wittgensteinian literature. Secondly, and more importantly, my difficulty arises from the fact that I am in agreement with what Suresh Chandra says at most of the places in his review.

My aim in writing this book was to 'present a concise exposition of the later Wittgenstein's philosophy of mind'. The book was not 'intended to be an exhaustive account of Wittgenstein's philosophy of mind'. It was intended 'to be only an analytic study of its foundations'. It tried to do so by employing 'clear and simple conceptual apparatus' to explain the complex and abstruse notions used in the philosophy of mind. Its more general aim was to dispel the belief of some readers of Wittgenstein that he had a destructive mind and that reading his works 'has a rather corrosive effect in the sense that at the end of a reading of his ideas one is left with nothing positive'. I had most clearly and unambiguously formulated my program for the book when I stated on p. 5: 'In the present study, I have tried to develop an argument to show, quite contrary to the popular, and indeed influential, philosophical opinion, that there is no epistemological problem of mind whatever, and that the widespread philosophical scepticism with respect to our knowledge of minds has no foundations at all.' To carry out my programme, as stated in the very next sentence, 'As a heuristic principle. . . I have throughout followed what one might call a Wittgensteinian dictum of *placing everything as it is undistorted before us*—to create complete clarity which leads to the complete disappearance of the philosophical problems.' Suresh Chandra in his review seems to have from my viewpoint very significantly, missed this point relating to the general purpose of the book. I may even say that he has missed the wood for the trees.

Generally, I would have wholeheartedly supported Suresh Chandra in choosing the way of the 'purist', rather than the way of the 'impurist' that I have actually chosen. But in the case of Wittgenstein one has to risk 'the danger of looking at Wittgenstein in the image of some other philosopher'. It has to be done in order to understand why Wittgenstein, when giving reasons for choosing to get the book published, in his preface to *Philosophical Investigations* written in January 1945, wrote:

> Up to a short time ago I had really given up the idea of publishing my work in my lifetime. It used, indeed, to be revived from time to time: mainly because I was obliged to learn that my results (which I had communicated in lectures, typescripts and discussions), variously misunderstood, more or less mangled or watered down, were in circulation. This strung my vanity and I had difficulty in quieting it.

One of the questions often raised during the lifetime of Wittgenstein and continues to be raised even today is: Who were the scholars whom Wittgenstein had in mind when he wrote this? And in what ways were they misinterpreting or distorting his ideas? Wittgenstein, when asked how many people he thought understood his philosophy, replied: 'Two—and one of them is Gilbert Ryle'. (Ray Monk, *The Duty of Genius*, Vintage, 1991, p. 436) So much about Suresh Chandra's charge of my being an 'impurist' in trying to see Wittgenstein through the eyes of Gilbert Ryle and merging the 'identity' of Wittgenstein 'into the identity of so many British philosophers'.

It is true that the term 'category mistake' is, what Suresh Chandra calls, 'Ryle's jargon'. But he is certainly wrong when he concludes that 'Wittgenstein would certainly reject this jargon'. Neither Wittgenstein nor I would say that the sentence 'vowel *e* is yellow' exhibits a category mistake. I am fully aware that a category mistake 'arises when things or facts of one kind are presented as if they belonged to another. Someone would make a category mistake if after being shown all the battalions and regiments she wished to be shown the army. . . Thinking of beliefs as in the head, or numbers as large spatial objects, or God as a person, or time as flowing, may each be making category mistake' (Simon Blackburn, *The Oxford Dictionary of Philosophy*, OUP, 1994). Suresh Chandra's conclusion 'To Vohra' the sentence 'vowel *e* is yellow' 'exhibits a category mistake' has apparently arisen from his confusing between the notion of a category mistake and Wittgenstein's distinction between 'primary' and 'secondary' sense of a word. The confusion is bound to be there when an exemplar is taken out of its context. Wittgenstein has used the sentence as an example for making a distinction between the primary and the secondary sense of a word. For Wittgenstein 'It is only if the word has a primary sense for you that you use it in the secondary one'. He took the example of the sentence 'for me vowel *e* is yellow' to demonstrate the

distinction between the secondary sense of a word and its 'metaphorical' sense. For him 'the secondary sense is not a "metaphorical" sense' (*PI*, II, p. 216). One can explain this distinction with the help of the use of the term 'blue' in the following propositions:

(1) Blue is the colour of the sky.
(2) He wore a faded blue shirt that matched his pale blue eyes.
(3) He's in a blue funk about his job interview.
(4) One should invest in a blue chip company.
(5) He is a blue blooded fellow.
(6) Blue collared jobs are not as well paid as the white collared ones.
(7) He is the blue eyed boy of his mother.
(8) His type of humour is a bit too blue for my taste.
(9) Blue films are meant for adults only.
(10) He's been feeling blue since he failed his exams.

In the above examples the word blue in the first two sentences is used in a primary sense, but in the rest of them it is used in the secondary sense. The primary sense of a term, for example of the term blue, cannot be conveyed in any way other than by means of the idea that is intended to be conveyed. Whereas the secondary sense of a term can be conveyed with the help of various other ideas and in several ways. Wittgenstein had the primary sense in mind when taking the example of 'yellow' he said: 'I could not express what I want to say in any other way than by means of the idea yellow' (*PI*, II, p. 216). To confuse the primary sense of a term with its secondary sense may amount to committing a logical howler, a philosophical blunder, but it certainly is not what Ryle has called a category mistake. For him 'a category mistake. . . represents the facts. . . as if they belonged to one logical type or category (or range of types or categories), when they actually belong to another' (*The Concept of Mind*, Hutchinson, 1949). One who commits a category mistake, therefore, is not the one who uses a term either in its secondary sense or in its metaphorical sense. It is commited by those who are 'perfectly competent to apply concepts, at least in the situations with which they are familiar, but are still liable in their abstract thinking to allocate those concepts to logical types to which they do not belong' (ibid).

According to Suresh Chandra a solipsist commits the blunder of treating language as 'necessarily private'. This of course is true. But from this how Suresh Chandra arrives at the conclusion 'Not only early even later Wittgenstein committed this blunder' is not clear, nor has he advanced any argument to support his contention. If Suresh Chandra is right then he is propounding a radical thesis, for, by now it is well established in the Wittgensteinian scholarship that Wittgenstein does not support the theory of privacy of language. Actually he vehemently opposes any theory which either directly or indirectly supports even the possibility of a private language.

Likewise, Suresh Chandra maintains that Wittgenstein rejects the theory that 'I' can be the name of a person primarily because 'it is not a referring expression at all'. He thinks that I am wrong in maintaining that the reason for Wittgenstein's rejection of the theory which upholds that 'I' is the name of a person is that it is primarily indexical. Quoting *PI*, 410 where Wittgenstein explicitly says that though indexicals like 'I', 'here', and 'this' are not names themselves, 'they are connected with names. Names are explained by means of them'. Suresh Chandra asks: 'How can they explain names if they are just index words?' This question reveals his ignorance about the nature of indexicals. An indexical is 'an expression whose reference on an occasion is dependent upon the context: either who utters it, or when and where it is uttered, or what object is pointed out at the time of its utterance'. It is primarily because of the fact that the index words have the power to refer to a person, a situation, a place, a time, etc. that they are able to explain among other things the names also. Wittgenstein had this role of indexicals in mind when in *PI*, 38 he said 'it is precisely characteristic of a name that it is defined by means of the demonstrative expression? "That is N" (or "That is called N"). But do we also give the definitions: "That is called 'this'", or "This is called 'this'"? We never do so. It is because "we call very different things 'names'; the word 'name' is used to characterize many different kinds of use of a word, related to one another in many different ways—but the kind of use that 'This' has is not among them" ' (*PI*, 38). Instead, indexicals like 'this', 'I', 'that', are used among other things to explain names. To say that 'I' or other indexicals are not 'referring expressions at all' as Suresh Chandra asserts, is to contravene an accepted usage, as well as to traverse their role in language. Suresh Chandra has to explain as to why he considers that 'I' 'is not a referring expression at all'.

It seems to me that Suresh Chandra has misunderstood the formula, namely, 'Observation=Having a sensation+Paying heed to it' that I have derived from my discussion on the logic of each of the concepts of sense-data, sensations and observation. According to him, 'If one attends to one's sensations, if he pays heed to them, then one may not pay any heed to the object that lies outside one's sensations, the object that awaits observation'. From this he concludes that 'Obviously Vohra does not mean that in order to observe something one has to attend to one's sensations'. The purpose of putting this equation was to highlight the distinction between the concept of observation, that is, the content of awareness on the other—a distinction that has too often been confused in the philosophical discussion. How can one observe something without even having a glimpse of it? I had clearly stated this on p. 12 of my book in the following words: 'There is a contradiction, a kind of logical oddity, involved in saying that someone was looking or peeping, at something without having a single glimpse of it, or that someone was listening to something without having even a single auditory sensation'. In fact at any

given instant of time, or situation we are constantly infested with countless visual, auditory, gustatory, olfactory, or many other kinds of organic sensations like pains, aches, itches, etc. But we are not aware of all of them. We are at a given instant aware only of that, or those sensations to which we pay heed. Sometimes we may be so lost in our thoughts, or so engaged in the work at hand that we may not be aware of any of them. The formula, therefore, highlights the important role that sensations play in observation. Even a cursory understanding of the formula shows that it implies that relevant sensations in proper measure are a necessary condition of observation. This is what Ryle also had in mind when he says very clearly 'disagreeable sensations do not help, they hinder perception. . . Sensations, in the sense, are not usually present when perception occurs... They are not *sine qua nons* of perception'. Obviously 'They' here refers to the 'disagreeable sensations'. One may remember that an equation, for example, in Chemistry holds good if and only if it is balanced. An unbalanced equation does not yield a formula. So that if any of the components of an equation are present in an improper measure the formula is not valid. An excess of a sensation would certainly make the formula invalid. Had Suresh Chandra paid heed to the beginning of the quotation from Ryle's 'Sensations' cited by me on p. 18 and analyzed it in its entirety, he would not have concluded 'Vohra sticks to Ryle's earlier views otherwise he would have failed to produce the formula that he produced'. Nor would he have averred 'Ryle would himself reject Vohra's formula'.

Suresh Chandra's confusions are worst confounded in his discussion of the second chapter of my book. His confusion arises from not recognizing the distinction between identity and individuation which I am at pains to explain. My saying that "the case of pains is not like physical objects" is to be seen in the context, which once again Suresh Chandra has missed. I assert the above statement after I have established that "the confusion about identity lies in the mistake of thinking that the same is same always, that the use of the word 'same' is governed by the same fixed rule irrespective of the context; whether we may be talking of 'coats', 'tables', 'pains', 'gaits' or 'sensations'". Suresh Chandra has committed the mistake to which I had pointed a little earlier in this reply. His "mistake lies in not seeing the 'same' must always be understood not in an abstract and pure sense, but together with some general term such as pain, or coat, and the criterion of identity in any particular case is determined by the general term involved". In the absence of a general term the talk of the 'same' or 'identical' also does not make sense. One may say that the general term is that which, as it were, prepares a place for the use of words like the 'same' and 'identical' in the language game, but then so strong is the lure of the surface grammar of words that even a thorough Wittgensteinian like Suresh Chandra is bewitched of his intelligence by it. It is because of this that he is unable to see that in his

argument to establish 'colour behaves in the same fashion as a chair'; 'colour and chair, stand and fall together', he has to use 'colour-patch' and not the word colour alone. No one shall dispute that there are two colour patches of the same or identical colour. But, how can one say that there are two colours when one sees only one colour, say, red in both the colour-patches. The 'genus' in both is the colour, say red, and the differentia is the location. Just as generically the toothache suffered by three different persons is the same though it is located in the mouths of three different persons. It is because of the generic identity of toothache in the three mouths that the dentist treats them analogously, i.e., either by drilling and filling, or by local application of gum paint, or by administering sedatives, or by extraction of tooth etc. Had Suresh Chandra delved into depth grammar of the word 'chair' and 'colour', he would not have argued 'if it is impossible for a chair here to be numerically the same as a chair over there, then it is also impossible for a colour here to be numerically the same as a colour over there'. Whereas we say that the colour in two places is the same, we only say that the chairs in two rooms are similar. They are called similar only if they are made of same material, have the same shape, same colour, etc. In the case of chairs we normally say that the two chairs are similar to one another, or that they are replicas but in the case of colours and pains we say the two are same, or exactly identical.

Strangely, Suresh Chandra has accused me of confusing 'the bundle theory' and 'the substance theory' of mind. His conclusion and accusation are once again the result of selectively choosing a statement that suits him to support his view and interpreting it out of its context. Had he read the statement along with the statements that follow it he would not have accused me of either 'proving that Hume was Descartes in a new guise' nor would he have been 'surprised' and 'shocked' by what he terms as my 'unusual interpretation of Hume's view'. The statements that follow it are:

> On the substance theory, the reflective observer was encouraged to think of himself, of his palpable bodily self, as inhabited by a ghost, or rather as being that ghost. And, the introspection theory was raised to tell us about the way in which the ghost is supposed to directly observe himself and his acts and states as contrasted with the way in which he observes the outside world through the mediacy of the body it inhabits. This basic assumption of the two theorists, however is wrong, for when we talk of a person's mind, we do not talk of a ghost in the body, or a self-luminous substance, or a theatre of special status incidents, but of certain ways in which some of the (psychic) incidents of his life are ordered.

From this it should be obvious to any discerning reader that Suresh Chandra has no cause to be either surprised or shocked because I am

neither giving Descartes 'birth with the thought form of Hume', nor giving a 'new name' to either 'bundle theory' or 'substance theory'.

I do not know how Suresh Chandra arrives at the conclusion 'By formal features Vohra simply means "psychological features"'. At no stage in the book either overtly or covertly have I said so. Neither can it be inferred from any thing that I have said in the book. Like Wittgenstein by 'formal features' I mean those features which are 'beyond being justified or unjustified; as it were, as something animal' (*OC*, 359). They form 'the basis of action, and therefore, naturally of thought' (*OC*, 411). They are the foundations on which the edifice of language games stands, and are in turn 'not founded' (*OC*, 253). They are the 'unmoving foundation' (*OC*, 211) which are so fundamental to all our activities, our form of life, that raising questions about them 'would seem to drag everything with it and plunge it into chaos' (*OC*, 81). That I do not exclude bodily features from the formal features is clear from the following argument advanced by me, on p. 109 of my book, to show that we neither require pure and exclusive bodily criterion, nor pure and exclusive memory criterion for establishing personal identity:

> . . . some character traits, and mannerism are more suited to one body than to the other. For example, in the prince-cobbler body interchange case the prince's body might include the sort of face that just could not express the cobbler's morose suspiciousness, the cobbler's is a face no expression of which could be taken for the fastidious arrogance of the prince. These 'clouds' are not just factual but logical. Such expression on these features might be unthinkable.

So, I cannot be said to have ignored the bodily features from being included in the formal features. The above should be evidence enough to show that I have neither 'introduced the Cartesian ghost into Wittegenstein's philosophy', nor forgotten 'not only my Wittgenstein but also my Ryle' as feared by Suresh Chandra.

In fact, I consider bodily features so important that in my discussion of personal identity I have time and again said that bodily criterion is crucial, though *not the only* criterion of asserting personal identity. Our criterion of personal identity has to have a proper mix of behavioural component and memory component in it. The context would determine what is to be called a proper mix. It shall not be in tune with Wittgenstein's spirit to lay down in advance the *exact* combination of bodily and memory component of the criteria in each case. So much in reply to Suresh Chandra's question 'Is there any doubt that Vohra has supported the theory of other minds that involves bodily identity as criterion of personal identity?'

One should not forget that even in the case of 'formal features' of our lives when a claim is made about their knowledge, it has to be justified by showing that one is in a position to make the knowledge claim in question. Wittgenstein was at pains to explain this in *On Certainty*. For the first person psychological statements to serve as evidence, or for these to function 'as the foundational claims in a theory of knowledge', all that is required of one is to show that the person making the assertion is 'normal'. Wittgenstein says in *On Certainty* 441:

> In a court of law the mere assurance 'I know. . .' on the part of a witness would convince no one. It must be shown that he was in a position to know. Even the assurance 'I know that that's a hand', said while someone looked at his own hand, would not be credible unless we knew the circumstances in which it was said. And if we do know them, it seems to be an assurance that the person speaking is normal in this respect.

My reason for saying that my having learnt English is good enough to answer some questions like 'How do I know that this colour is red?' was that in such cases 'we do not have to look for any other reason, nor is there one'. This I had upheld to be applicable to all those cases where circumstances are such 'as it would be absurd for us to express this without hesitation or reservation. In all such cases although what I say is based on sense experience it is not known by any sort of inference. I do not have any evidence for what I say, and I could not now think of any evidence against it'. Apparently these conditions are not applicable to the example of arsenic taken by Suresh Chandra. And in a court of law in order to establish one's claim one has to show that he was in a position to know that X is a murderer. Once one is conclusively able to prove that he was in a position to know, no further question is possible, and if it is asked it is redundant. So, there is no question of my extending carelessly and irresponsibly the argument to each and every case.

I must conclude by saying that I immensely enjoyed reading Suresh Chandra's comments. I am thankful to him for giving me an opportunity to reopen a book which I had closed long ago. There still remain a number of issues which though raised in the characteristic non-serious style of Suresh Chandra, need serious consideration. I assure Suresh Chandra that his effort will not go waste and I shall soon develop them.

University of Delhi and Indian Council of Philosophical Research ASHOK VOHRA

Focus

Dr Mukund Lath has drawn our attention to the following text from Udayana's *Pariśuddi*, where he self-consciously proclaims himself to be an '*ādhunika*' and '*abhinava*' in contrast to Jayanta, etc. whom he considers as *jara naiyāyika*. It will be interesting to find as to what exactly are the features which he considers as distinguishing him from the 'ancients' and which, in his view, make him as '*ādhunika*' naiyayika. He has also made a five-fold distinction in the relation between *vācya* and *vācaka* which needs to be paid closer attention by contemporary philosophers in the analytical tradition.

अत्रोपमानस्य फले विप्रतिपद्यमानान् प्रति साशङ्कं जरन्नैयायिकजयन्त प्रभृतीनां परिहारमाह–यद्यपीति। संज्ञाकर्म संज्ञाकरणम्। मानान्तरं शब्दो वा शब्दोर्थो वा लिङ्गभूतः। तत्र लिङ्गपक्षे, उपलक्षितः। शब्दादवगतेन गोसादृश्येन प्रतिपाद्यमानः पिण्डः समानासमानजातीयेभ्यो व्यवच्छिन्नः शक्यमवगन्तुमिति संबन्धः। शब्दपक्षे तु, गोसादृश्येनोपलक्षितो विशिष्टः शब्दादवगम्यमानः पिण्डः समानासमानजातीयेभ्यो व्यवच्छिन्नः शक्योऽवगन्तुम्, गुणकर्मादिवत् सादृश्यस्यापि समानासमानजातीयव्यवच्छेदकत्वादिति सुगमम्। गोसदृशो गवयशब्दवाच्य इति तावदस्तु। अथ गोसदृश एव क इत्याशङ्कानिवर्तनासामर्थ्यमतेः परिप्लवः। एनमेवार्थं कञ्चिदिन्यादिना दर्शितवान्। यत्रेति विषये। तर्हि विशिष्टैव प्रतिपत्तिर्वाक्यादस्तु इत्यत आह–न चेति। प्रत्यक्षमात्रद्भविष्यतीत्यत आ–न चेति।।

एतेन वाक्यप्रत्यक्षसमाहारोऽपिनिरस्तः। स हि प्रमाणसमाहारो वा स्यात्, फलसमाहारो वा? आद्येऽपि प्रमाणत्वे सति समाहारः, समाहृतयोर्वा प्रामाण्यमिति? न तावत् प्रथमः, फलानेकन्वे सति समाहारानुपपत्ते; तस्यैकफलं प्रति मरस्पराधिपत्यरूपत्वात्। नापि द्वितीयः, वाक्यप्रत्यक्षयोर्भिन्नकालत्वात्। वाक्यतदर्थयोः समृतिद्वारोपनीतावपि गवयपिण्डसंबद्धेः। फलसमाहारे तु तदन्तर्भविऽनुमानशब्दयोरपि प्रत्यक्षत्वप्रसङ्गः। तत् किं तत्फलस्य तत्प्रमाणबहिर्भाव एव, अन्तर्भविवा कियती सीमा? तत्तदसाधरणेन्द्रियादिसाहित्यम्। अस्ति तर्हि सादृश्यादिज्ञानकालेऽपि विस्फारितस्य चक्षुषो व्यापारः? न, तस्मिन् सति तस्यानुपयोगात्। उपलब्ध गोसादृश्य-विशिष्टगवयपिण्डस्य वाक्यस्मृतिमतः कालान्तरेऽप्यनुसन्धानबलात् समयपरिच्छेदोपपत्तिरिति। तदेतत् जरन्नैयायिकमतमास्कन्दति–नन्विति। पूर्वे हि यथा गौस्तथेत्येव वाक्यं गवयपदेनाप्रतीतसमयेनापि प्रयोगमात्रोपयोगिना सहितं संबन्धप्रतिपत्तिहेतुः, प्रतिसंबन्धिनस्तु शब्दाद् वोपलक्ष्णाद् वोपस्थापनमित्याहुः। आधुनिकास्तु प्रतिसंबन्ध्युपस्थितिः पूर्ववदेव, समयपरिच्छेदस्तु प्रयोगानुमानादित्याहुः।।

अवाचकप्रयोगस्य वाक्यानङ्गतयेति तद्विरहिणश्च संबन्ध्युपलक्षण-पर्यवसानादिति मन्यमाना एतदेव प्रयोगमारोप्य दर्शयन्ति–यो हीति। एतेनैव परिप्लवो निरस्तः। यदि हि संबन्धो विदितः, कः परिप्लवार्थः? अथन विदितः, किं तदुपवर्णनेन? विशेषनिष्ठतया न विदित इति

तत्रकाङ्क्षनुवृत्तेरेवमुच्यत इति चेत्–तदप्ययुक्तम्। न हि विशेषनिष्ठः समयः, तथा सत्यशक्यग्रहत्वात्, अपि तु किञ्चित् निमित्तीकृत्य देशकालादिविशेषविनाकृतव्यक्तिमात्रनिष्ठः। तदेव निमित्तं नावगतमिति चेत्? अवगतिरेव तर्हि संबन्धस्य संबन्धिनोऽनवगमादिति किं परिप्लुतं प्रमातुः? न चैष परमार्थः, सादृश्यविशिष्टपिण्डोपक्षेपे निमित्तस्याप्युपक्षेमात्। अन्यथा पिण्डोपक्षेपेऽपि निमित्तानुपक्षेपे मनोधर्मादिशब्दानां जातिशब्दत्वं न स्यात्, अवाचकत्वमेव व स्यात्, तज्जातीयानां नित्यपरोक्षत्वादिति।।

अत्रोच्यते अभिनवनैयायिकैः, पञ्चधा हि वाच्यवाचकभावव्यवस्था। निमित्तेपलक्षणरहिते शृङ्गग्राहिकया चैत्रमैत्रवत्। निर्निमित्तेऽन्तर्भूतोपलक्षणे परिभाषणिकया आकाशादिशब्दवत्। सनिमित्ते तटस्थोपलक्षण एवंप्रकारिकया पृथिव्यादिवत्। विशेषवन्निमित्ते निमित्तसङ्कोचनिकया स्वर्गादिवदिति। तत्र प्रथमद्वितीय-योरसंभावितत्वादिह विदितवृत्तान्ततया तृतीयं प्रकारं निरस्य पञ्चमश्चतुर्थमुपयातीति चतुर्थव्यवस्थापयति–न तावदिति। अविदितगवयशब्दवृत्तान्तस्य तु यद् भविष्यति, तदनन्तरमेव वक्ष्यति। तदिदमुक्तमिति। परमार्थो पस्तुगतिः ततश्च तथाभूत एव पिण्डे सेबन्धो ग्राह्यो न तु सादृश्योपलक्षणोपनीते पिण्डमात्र इत्यर्थः।।

* * *

Sūtra 1.1.4 in the *Vaiśeṣika Sūtras* by Kaṇāda has generally been regarded as the ground for accepting the basic number of *padārthas* mentioned in this system. However, there is some strong evidence which casts doubt on this and raises the question as to whether this *Sūtra* formed an integral part of the original *Vaiśeṣika Sūtras*, as it does not occur in some of the important manuscripts ascribed to well-known thinkers belonging to this school. We publish below an extract from Wilhelm Halbfass's book *(On Being and What There Is,* State University of New York Press, 1992, p. 75) and would like to know what others say in this connection:

> As we have noted earlier, a complete list of all six categories as well as the term *padārtha* itself is found only in one single *sūtra* of questionable authenticity. In the version of the *Vaiśesika Sūtra* that forms the basis of Śaṅkaramiśra's commentary, *Upaskāra*, it appears as 1.1.4. It is missing, however, in the versions used by several apparently older commentaries, that is, Candrānanda's *Vṛtti*, Bhaṭṭavādīndra's *Vārtika*, and the anonymous commentary edited by Anantalal Thakur.
>
> What does this mean? Is it merely a textual problem, or does it have deeper historical significance? Does the absence of *Sūtra* 1.1.4 in some versions of the text reflect a stage of development at which the list of six categories (which is taken for granted not only in Vātsyāyana's *Nyāyabhāṣya* but also by Caraka and Āryadeva) had not been finalized? The *sūtras* that follow 1.1.4 provide a systematic and coherent presentation of substances, qualities, and motions. Universals, particularities, and inherence are introduced later, in a somewhat casual and erratic fashion. Could this be an indication that there was an earlier version of the *Vaiśeṣika* that dealt only with the first three of Praśastapāda's six categories?

* * *

The *Mīmāṁsā Sūtras* are generally ascribed to Jaimini. However, there is almost incontrovertible evidence that such an ascription cannot be sustained in the face of the counter-evidence in the *Mīmāṁsā Sūtras* themselves. We publish below an extract in this connection from an early study of Vedānta* by Hajime Nakamura, a well-known Japanese scholar on Indian philosophy and thought:

> According to tradition, Jaimini is the author of the *Mīmāṁsā Sūtra*, and this has been accepted throughout India. However, the *Mīmāṁsā Sūtra* handed down to the present time was not directly written by Jaimini himself. In this *Sūtra*, the name of a scholar called Jaimini is mentioned five times, and for the author of the *Sūtra* to give his own name within a *sūtra* composed by himself would be unusual. Not only that, in the *Mīmāṁsā Sūtra* VI. 3.4, Jaimini is referred to as the supporter of the opponents and is later rejected. Consequently, the compiler of the *Sūtra* is clearly different from Jaimini, and while they respected and relied upon him as an authoritative thinker, he did not follow Jaimini's theories in all respects, but maintained a critical attitude towards them. It is not clear to this day just who the author of this *Sūtra* is. Perhaps it was compiled as a kind of text book on ritualistic *Mīmāṁsā* learning by the *Mīmāṁsā* scholars after the time of Jaimini. Since in the extant *Mīmāṁsā Sūtra* the names of a great many others besides Jaimini appear, it might be a compendium of their theories. If so, one can suppose that quite a long period of time was necessary for it to assume such a form.

* *A History of Early Vedānta Philosophy*, Hajime Nakamura, Motilal Banarsidass, 1983, pp. 390-91.

Agenda for Research

Śaṁkara, in his bhāṣya on the *Brahma Sūtras* criticized almost all the schools of philosophy known in India during his time. As most of these schools continued to flourish after Śaṁkara, it is important to find their replies to his criticism of their position. This would also show the extent of their awareness of Śaṁkara's criticism as well as the period when they took serious note about what he had said about their positions. The Buddhists, for example, continued to flourish till almost 1200 A.D., that is, for almost five hundred years after Śaṁkara. Yet, except for Jñāñaśri Mitra (1025 A.D.) who had written a work entitled *Advaitabindu* which presumably discusses Śaṁkara's argument against *vijñānavāda*, no one else is generally known in the tradition for having done so. Similarly, it would be interesting to find when the Jains, the Sāṁkhyans or the Vaiśeṣikas took note of Śaṁkara's criticism. This would give us some idea of the gradual penetration of the awareness of Śaṁkara's thought among the different philosophical schools of India and their reactions to what he had said. It will also be interesting to find whether the Advaitic thinkers took note of these objections and replied to them and, if so, when? In fact, while the debate between the Advaitins and the non-Advaitic Vedāntins is fairly well known, as is also the debate between the Advaitins and the Naiyāyikas after Śaṁkara Miśra, the debate between the Advaitins and other schools of Indian philosophy has been given very little attention up till now. It is time that this gap in our information is filled and the different dimensions of the debate compared, evaluated and critically examined.

Notes and Queries

It is usually held that the Naiyāyīka cannot accept a sentence such as '*ghaṭo-ghaṭaḥ*' as meaningful in a system for, according to him, any sentence to be meaningful must give some new knowledge. However, we have received the following statement from Professor V.N. Jha of Pune University arguing that the Naiyāyīka will have to accept the sentence '*ghaṭo-ghaṭaḥ*' as meaningful, if he wants to stick to his definition of *anyonyābhāva* as the latter entails the former:

The Navya Nyāya provides the definition of *anyonyābhāva* or *bheda* as follows:

तादात्म्य-सम्बन्धवच्छिन्न-
प्रतियोगिताक-अभाव:

'A mutual absence is that absence the contra-positive of which is delimited by the relation of identity.'

The example may be paraphrased as either

(1) पट प्रतियोगिक भेदवान्-घट:
or (2) घट प्रतियोगिक-भेदवान्-पट:

Let us expand either of them:

घट प्रतियोगिक-भेदवान् पट:
=घटनिष्ट प्रतियोगिता-निरूपक भेदवान् पट:
=तदात्म्य (=अभेद)-
सम्बन्धावच्छिन्न-घटानिष्ठ प्रति-योगिता-निरूपक-भेदवान् पट:

implies that

घट घट:, घट: न पट:
पट: पट:, पट: न पट:

Unless this is accepted the definition of भेद cannot be justified according to Nyāya. As a matter of fact, the tradition says that घट: अभेदेन घटे अस्ति, अभेदेन् पटे नास्ति.

That is what I mean when I said *ghaṭo-ghaṭaḥ* has to be accepted by Navya Nyāya.

* * *

Nāgārjuna, in the various works attributed to him, criticises the possibility of any *pramāṇa* being given to establish anything because, according to his analyses, it is intrinsically incapable of doing do. In case it is correct, one would expect that there would have been refutation of his position

in the Buddhist tradition itself so that any Buddhist *pramāṇaśāstra* could be constructed. But, little is known of such refutation even though Dignāga wrote his *Pramāṇasamuccaya* and Dharmakīrti a *vārtik* on it. It would be interesting, therefore, to find the post criticism of Nāgārjuna and compare them with those made in the Nyāya tradition.

* * *

We have received the following clarification regarding the distinction between *Logos* and *nous* in Greek philosophy from Professor Richard Sorabji, King's College, London:

> In Greek thought, *Logos* is contrasted with *nous* from Aristotle onwards. The former refers to a step by step reasoning process which leads to an understanding while the latter (*nous*) does not need to do so as it has already arrived. *Nous* is sometimes thought of as intuitive. In the Neo-Platonist tradition it is considered as timeless. Episteme in Aristotle believes *scientific* understanding. Aristotle means that episteme passes from the premises (often definitions) which are initially grasped by *nous* from which inferences are drawn through *Logos* usually by syllogistic reasoning to an understanding of the further attributes of the thing defined. For example, it passes from the definition of triangle to seeing that it follows from the definition, with other premises, that interior angles will be equal to two right angles.

Book Reviews

SOM RAJ GUPTA, *The Word Speaks to the Faustian Man,* Volume II, Motilal Banarsidass, Delhi, 1995, pp. 535, Rs. 550.00

The volume comprises translation and interpretation of the *Mundaka Upaniṣad* with the *Śaṁkara Bhāsya* and (2) the *Māndūkya Upaniṣad* with the *Śaṁkara Bhāsya* and *Gouḍapāda Karikas.* This is part fulfillment of what the series 'The Word Speaks to the Faustian Man' pledges, namely, translation and interpretation of the *Prasthanathrayi.* At the outset, there is something striking and singular about the whole enterprise—its method, plan and purpose—which calls for attention. The Preface tells us, this is 'not a mere work of Indology but an explicit critique of modern civilization and its myths and ideologies'. Obviously, the purpose of Sanskrit learning has to do with emphasis on philosophical and cultural traditions in regard to substance, method of analysis and expression. Instead, the focus is to be shifted firmly on to the practical aspect—to what those ancient conceptual formulations and Upaniṣadic wisdom might mean in experiential terms. At the same time the 'Bhāsya' style of discourse is integrated into an amazing mutation, harmonizing the ancient with the modern idiom of articulation. Yet the two polarities remain clear without any blurring or equivocation.

There has always been a feeling that the use of western philosophical idiom to bring out the sense of the Indian thought proves frustrating, because of its being either inadequate or actually distorting. Here one is delighted to see the western philosophical idiom being freely used with the least strain to convey Upaniṣadic insights free of distortion or loss of nuances. (This volume could supply some truly exquisite role-models of that kind.)

The sub-title to the series informs us this translation and interpretation of the *Prasthānatrayi* is being offered 'for the participation of the contemporary man'. So that 'participation' (प्रतीति) and 'contemporaneity' are the guiding sign-posts. This closeness to the stuff of lived experience is a pervasive influence and invests the discussion in various ways. In the Preface, for instance, one lights upon a brilliant little vignette imaging the Upaniṣadic man as a living entity relevant even for the contemporary world. The creative centre of that civilization wasn't either village or town but *aranyaka,* forest. Man lived in a cosmos with gods as his neighbours and beasts as his kin. This 'purusha' nursed in that sort of *tapovana* absolutely at a peace with air, earth and sky would 'fashion the face and soul of civilization—this humble nonentity'.

True, that forest civilization came to be ruined as caste divorced from the 'āśramas' and the renunciatory spirit turned into evil and the

civilization got perverted. But the modern man who has built his tentative system of civilized existence on rationalist basis (what he considers rational) has need to turn to that nonentity of the erstwhile forest-polis civilization for enrichment, even for survival. In other words, the Upaniṣadic man is very much with us—alive and operative. Hence the author's exhortation in the Preface: 'Let us, even at this belated hour, learn to expand the bounds of our civilization. For our home is not the human *polis* but the earth and the sky. And death. If we do not wish to commit collective suicide, we have to learn to find a place in our midst for these others. The Faustian man must stop his prattle and wait in patience for the peal of the word.'

That should explain a good deal about this book's aims and the strategy followed. The commitment to be bringing everything as far as possible to the test of actual experience informs the discussion throughout. That accounts for one or two salient features of the series. There is a natural bent towards using highlights, great inspired moments, of the poetic art to illustrate or throw new light upon some abstruse metaphysical phenomenon—to bring it within the ambit of comprehension. Response to such a piece of the literary art offers direct and immediate access to the experience involved, cutting through abstractions, excesses of conceptualization and discourse. Needless to say, such a practice draws upon the profound relationship—amounting to a rich confluence that takes place between the two, between the metaphysical and the poetic, between poetry and religion.

Thus, crucial passages from Shakespeare's *Macbeth, King Lear* and *The Tempest,* Wordsworth's *The Prelude,* Coleridge's *Dejection: An Ode,* Tolstoy's *War and Peace,* Dostoevsky's *Letters from the Underground,* and so on, not only bring illumination by focussing a strong shaft of light on the point at issue but help carry conviction by awakening and engaging our imaginative power. In fact, the title 'The Word Speaks to the Faustian man' itself has a direct bearing on literature. It evokes the Faustus of legend, the tragical symbol of modern man's quest for limitless knowledge, pleasure, skill and power with the inevitable cost he has to pay up for it. Echoes of the magnificent *finale* of Marlowe's play with his Doctor Faustus poised between terrestrial time and eternal damnation are in our ears. ('See, see, Christ's blood streams in the firmament. One drop would save my soul half a drop. . .') So also, Goethe's final vision of *his* Faust giving himself up to the task of reclaiming the swamp-land for humanity would be haunting us to keep company with this book's main thrust and argument.

This link with literature is one of the delectable features of the entire discussion. It means, much more than merely a delectable feature superadded. It means, such peaks of the literary art the Faustian man can share with his Upaniṣadic counterpart. This rich vein of pure gold in western literature serves as a common ground for the two to share the poet's

insights—feel and be in unison; so that when the word peals, the Faustian man can respond with an answering celerity and eagerness!

Another allied feature helps us place the author's enterprise in its proper 'activist' perspective. In the Indian tradition there is always at the back of every quest for spiritual knowledge an assumption—voiced or unvoiced—that the realization of the truth of things demands as a precondition a certain purity in the seeker which enables him to merit it, as it were. A certain condition of mind and soul is to be reached first such as would facilitate his perception of truth, entitle him to it, give him the *adhikāra* for it, as it were. Intellectual effort alone, by itself, is not adequate. The total man has to be readied to receive the truth, may be, through the practice of *yoga.* It is therefore appropriate that an authentic description of the essentials of the yogic exercise have been included in this volume. More valuable still is the careful account of some significant scenes from the life of the author's Guru, Sri Mangatram, are provided—what might aid the reader's effort to image for himself the condition that is regarded as *Siddhi.*

What is central to this fresh approach is that it highlights and clarifies certain recondite issues that hamper our understanding of the ancient text. If not so clarified or resolved, these become hurdles in our getting to the core of the text. That's how the author puts things in a modern perspective. He would step in from time to time to explain some peculiar aspect of ancient thought or articulation in modern terms in the contemporary idiom . . . what might otherwise seem excessive, trivial, or downright perverse. Such bringing to mind of the underlying context of thought proves crucial.

Thus at the beginning of his commentary on the *Mundaka* he makes clear what is implicit in the somewhat enigmatic nuance of 'The Word' that occurs in the title, and the way it is to be received. That's a basic issue for the whole enterprise. *Śruti,* the Vedic word, carried significances similar to those of 'logos' in St. John. Here too, as the commentator puts it, 'In the beginning the Word was with the creator and the Word was the creator . . . Between the Word and the creator there was no gap, as there was none between the creator and the creation.' The Vedic word is the revelatory word which the seers see directly, it being reality itself. The Vedic word is not a multivalent word that could yield different meanings to different dispositions. The only disposition in which it can awaken a response is that of a passive innocence and self-submission to it. The Vedas reveal themselves to one who will hear them in passive innocence, not to those for whom words are a means. For the scholar, the Upaniṣads are referential; to the submissive soul they are reality itself.

But when the word was transmitted to Atharva and to men, it gets transformed into the lack that resides at the heart of human language. 'Being' has deserted the word which has become referential. So that bedevilled with multivalence, interpretations, discourses, we can now

never quite bridge the gap between word and reality. An authentic grasp of this attitude to language would make all the difference to our understanding of the Upaniṣad. It is indeed a singular feature of this commentary that it throughout offers many such illuminating insights outside the immediate scope of the textual exegesis of the Vedic word.

The commentary clarifies another feature of discourse allied to this one. It is something of crucial importance to Upaniṣadic thought and of continued relevance to textual exegesis. That feature is the use of analogy in the form of simile and illustration. A knowledge of the manner an Upaniṣadic simile functions can help elicit an appropriate response. First of all, how is it that the Upaniṣads seem to prefer simile and illustration to philosophic argument? Here the commentator elaborates the point of the basic premise: The whole of reality can neither be grasped by the human intellect nor can it be expressed fully through the limited language of men. The Upaniṣads are based on intuitive human knowledge, experience and revelation. Truth is not imagined or rationally determined; it is intuitively experienced. So the method adopted by the Upaniṣads is to convey knowledge pertaining to the Self through examples, illustrations and similes.

Again, as the commentator points out, the arguments of philosophers are anthropomorphic and culture-based. Man is so made that he asks questions and answers them within a theoretical framework. If the Upaniṣads resort to simile to reveal the truth, how does the simile succeed where the argument of the philosopher or the theory-laden observation of the scientist fails? How does it break the anthropomorphic confines which keep man shut off from truth and reality? The Upaniṣad does not directly express or explain. It functions only through the pregnant suggestiveness which the simile evokes. So that the deeply disillusioned soul of the reader approaching the Upaniṣad in an appropriate mental posture has only the simile before him. He has no concept to argue about, understand, or to get clarified, but just the concrete simile that stands out becoming the sole avenue to truth, as is the case, for instance with those celebrated verses of the *Mundaka* figuring the spider (यथोर्णनाभिः सृजतेगृह्-णतेच) or sparks from blazing fire (यथा सुदीप्तात् पावकात्• • • •) or the two mythic birds (द्वासुपर्णा सयुजा सखाया • • •) or the coursing rivers eager to merge into the ocean. (यथानद्यः स्पन्दमानाः • • •) The simile points to a fact before the reader, the fact which he can readily perceive and experience. Let him, the Upaniṣad suggests, look at it without pre-suppositions, without theories. The observation has to be innocent, 'with no cognitive axe to grind'. He has to be so wholly absorbed in concrete observation of the object, the act, the event, as to achieve an empathetic oneness with it. So close is the identification. Let him be so absorbed in the act of the spider, says the commentator, that he becomes the spider and spins out and withdraws the thread. All of which effectively dissolves or melts the theoretical framework of anthropological thinking.

The similes the Upaniṣads use are invitation to man to get absorbed in the object before him. To the man of argument or of theory-laden observation this invitation will mean bad philosophy and primitive thinking. But to the 'participatory' soul who can spin out and spin in with the spider the simile is the sacred word—the only way to the truth. To the 'cogito', detached and withdrawn, an analogy is just an instance; for a 'participator' it is self-discovery, his truth and being. The commentator visualizes a deeper basis for this kind of phenomenon. He points out that Vedāntic insight has discovered an interiority in man which is impersonal, sense-free, and action free. The singular manner the simile functions in an Upaniṣadic text owes to such an interiority and in effect testifies to it.

This feature of scriptural discourse has a wider bearing even outside Vedāntic thought. In the Buddhist text, *Milindpanha*, for instance, we find the simile being credited with the same kind of power. Whenever King Milinda (Greek General Menander who had embraced Buddhism) finds Nagasena's propositions or arguments unconvincing he exclaims, 'How's that?' and asks for a simile. Thus we see him continually demanding: 'Give me a simile; give me an illustration; give me a comparison. . .' and so on, step by step, for a better assurance of truth that the pondering of an analogy provides.

I have dealt with this question of the simile at some length hoping it might serve as a sampling to indicate the nature and value of the intellectual stimulation one is offered here. In general, what is aimed at is to help the modern reader to clear the hurdles that are likely to frustrate his efforts to enter into the spirit of the ancient word that the Upaniṣads represent. Mainly, the attempt is to free him of the ingrained Cartesian-Newtonian habit of mind, and look upon the Upaniṣad freshly in its own terms as a practical, realizable, living experience to-day for the human being who is a 'participator'. By what means is the Upaniṣadic truth brought close to modern man's feeling? A quick survey of a few such major items is all one could touch upon here.

For instance, the commentator finds occasion for it when, early on, the *Mundaka* speaks of the Vedic hymns being revealed to sages, to genuine individuals. The genuine individual is marked off from the man of the communal herd. The sound of waters stills his thoughts and the flight of white cranes give him their airy upleap. For him the silent pauses between words become a harmony, a *nāda*. To such a one revelation may come in words. Lest this sound absurd and sceptics scoff at it, the author narrates his personal experience of an incident from his biography of his preceptor, Sri Mangatram. He describes how once, 'eyes closed and oblivious of the world', his Master sang verses for hours on end in the colloquial traditional language of saints. He sinks into *samādhi* and the word peals out into words. To his inquisitive disciples he explains that this is peal of *prakriti*, the prime doer. 'The voice of nature arises to mingle into its air,

its breath. . . . Of God-intoxication is this flow, absolutely spontaneous and unmeditated. No thinking is involved in it, no conscious *making*.'

To the query of a disciple, 'The divine speech flows out mostly in Hindi, whereas you never learnt this language. You learnt only Urdu.' The Master says, he himself was surprised, but that the *vani* is in accord with the needs of the common people. Sometimes, as in the case of the Cosmic *Purusha* (पुरूष) it is not that easy bringing the Upaniṣadic experience within modern man's reach. To undo anthropomorphic implications the Upaniṣad denies any form to *the Person* ('Shining and formless is *the Person*'). In any case, Upaniṣads can impart higher knowledge only to a tranquil, innocent and passive soul. This is a condition and need which receives strong emphasis through out. And since the *Mundaka* itself describes the *Purusha* in terms of yogic postures and processes (अग्निर्मूर्धा etc.) the commentator contents himself with a fairly detailed explication of the functioning of the *Kundalini* Yoga. . . And Yoga is the way that tends to the realization of the *Purusha*.

Sometimes—as in the case of the celebrated pair of birds on the tree (द्वासुपर्णा सयुजा सखाया)—the commentary brings home the Upaniṣadic message direct to the contemporary man, even citing a familiar illuminating piece of Shakespeare in support. The tree of life is a brittle tree (the body). To it have come to roost both mortal soul and immortal God, constant companions, since where man is, there is also God. Man eats the fruits of that tree whereas God tastes them not. He merely looks on. The commentator clarifies and elaborates the little vignette. Man is a *telos*-oriented creature. But pleasure (in eating the fruits of that tree) may tempt him away from the *telos*. That's the result of his folly in taking this world and its business as the true Reality. And then does the commentator quote, as a counter-blast and corrective, the famous speech of Shakespeare's Prospero ending with the words: 'We are such stuff/as dreams are made on; and our little life/Is rounded with a sleep.'

One or two instances of this felicitous use of literature to engage contemporary interest may call for special note: For a participant soul even the gleam of a withered leaf that we trample under foot, unconcerned, may have the power to redeem us from hopelessness and despair. The illustrative artifact cited is from Dostoevsky's *Letters from the Underground*. A man is ordered to be shot for some little 'crime' against the Czar. He is tied to the post along with co-brothers of the 'crime'. The sun going down in the West sends but a small ray which lights up the metal cross on the top of the tranquil high-tower of the Church, in front. That gleaming cross atop the Church-tower arrests the convict's attention. What was there in that radiant metal gleam that could undo all his life-concerns? That moment of God-vision is one of ecstasy that we mortals call *Sacchidānanda*, the realization of *samata*, the moment of becoming the self.

Similarly when the issue under consideration is whether the sage, a

knower of Brahman, is worthy of worship, the commentator describes what he saw with his own eyes—Sri Mangatram, his preceptor, sitting in *samādhi*. 'No familiar human expression was visible in what had the appearance of a human face. . . It spoke of the beyond, of dimensions utterly unfamiliar to our thought or imagination. Moments arose when the body seemed to melt into a peaceful nothingness. In that nothingness seemed to lie the secret of our life, the meaning of our destiny.'

Lastly, one of the finest instances of an illustrative use of literature refers to *King Lear*. It helps elucidate a certain aspect of the general human condition. Our life is future-oriented. Tied to the past, entangled in the present, it must move on transcending both. Tensions between what is lost and what found generate storms. For a glimpse of these, the commentary invites attention to 'the great tragic masterpiece of the western world—Shakespeare's *King Lear*.' It is also a most extensive such reference. The entire action of the play is sensitively rendered, bringing out the significance of each stage of the change taking place in Lear. The impotent wrath, the heath-scenes, the onset of madness, the 'God's spies' speech, and all—the whole sequence is beautifully sketched. At the end, the commentator makes his observations on the state of mind and soul Lear has arrived at finally. These might strike one as somewhat at variance with the general opinion of discriminating readers and viewers of the play.

He makes the point that, as the 'God's spies' speech indicates Lear wishes to detach himself from the world-context: 'He would', says the commentator, 'live like a God's spy, some kind of an immortal; his rejection of the world-context is contaminated by almost the same sin which made him blind to human contextuality, made him believe that he was an *absolutus*, a being independent of all constitutive relationships . . . Can he, a creature of flesh, behave as if he were a detached, unconcerned, God who could live as a bemused witness of the foolish ways of men? Lear's sacrifice is incomplete, hopelessly incomplete. He has yet to come to terms with mortality. . . He would play the immortal in a world of make-believe. . .' And so, when they hang Cordelia because she and Lear stand in their way, Lear breaks down: 'And my poor fool is hanged!. . .

Look there, look there (dies)'

The point is well taken. In fact, it is an imaginative affirmation of a principle that distinguishes this work as a whole, and is implicit in earlier remarks like: 'Freedom from the world does not mean detachment from it; it means becoming one with it.' However, in the interest of maintaining a balanced perspective one or two considerations have to be allowed their due weight and plausibility.

Shakespeare was after all working within the framework of western tragedy, for one thing. And again, the kind of detachment Lear arrives at could be looked upon as a sort of pre-condition or a step in the

direction of the goal which has been more integral to *our* culture than to the western. . .

What is said here so far about the commentator's effort to engage contemporary interest in the Upaniṣadic experience through various means applies equally to the *Mundaka* and the *Māndūkya*. The commentary on the latter is very much more extensive, dealing, as it does, with the entire corpus of the *Gouḍapāda Karikas* on the *Māndukya*, numbering some 215 of them, in addition to the *Śaṁkara Bhāsya*. Moreover, though the *Māndūkya* itself is short the *Śaṁkara Bhāsya* and more often the *Gouḍapāda Karikas* give rise to discussions on metaphysical aspects which tend to be spelt out at great length. To mention one such theme rather at random, Śaṁkara's critique of the Buddhist idealistic position that there are no external phenomena receives considerable attention. There is some detailed argumentation on this issue. It is a brilliant discussion setting out the exact position Śaṁkara and Gouḍapāda hold—what could be 'genuine idealism'. (The cogency and comprehensive reach of what is convincingly argued here constitutes a finest achievement of this book.)

Gouḍapāda and Śaṁkara *appear to* accept both the Realist and the Idealist positions—as though hunting with the hounds and running with the hare. What resolves the clash of polarities is a kind of existential approach to actual experience. Both Śaṁkara and Gouḍapāda would point out, 'when you flow out to the other and the other to you, when you live duality, then both you and the object will turn into something other than you and other than the object. The Vedas call that something *Atman*, the Self. Others may call it, if they like, *shoonya* or liberation. To find the object as not the other, to love it, to flow out to it, is to die into it. That is the genuine truth of genuine Idealism.' Whatever the mind experiences is an unreal appearance. When the thing is really seen, it gets transformed into the Self. Thus two things this lengthy critique establishes: (1) that the Buddhist idealistic position of denying existence of external phenomena is untenable and (2) that there is essentially no difference between Śaṁkara's and Gouḍapāda's positions.

Western literary classics cited continue to illuminate and strengthen an argument: A passage from Wordsworth's *Prelude* brings home to us that the fundamental truth of life is inter-subjectivity, not subjectivity. The blessed babe nursed in the Mother's arms sinks to sleep rocked on her breast. he is an inmate of this active universe. 'Along his infant veins are interfused/Gravitation and the filial bond/Of nature that connects him with the world', since he is 'one who with his soul drinks in the feeling of his Mother's eyes'. And this early infusion of inter-subjectivity must last him all his life!

Similarly, literature's aid is sought when the question addressed is: 'Is there such a thing as non-cognitive awareness?' Cognitions are other-oriented and not self-revelatory. A non-cognitive awareness reveals these cognitions. Such an awareness shows itself when one steps out of one's

senses, thoughts and feelings in deep pain and anguish. The fitting illustration cited here is Coleridge's poem, *Dejection: An Ode*. The man is almost undone but there is a consciousness in him that notices the entire state of disruption with integrity. The 'person' is on the point of elimination when pure consciousness reveals the presence of the 'witness'. Similar is the condition of Prince Andrey from Tolstoy's *War and Peace* in the passage cited next. The Prince opens his eyes after he had lain on the battle-field unconscious, grievously wounded. He sees the lofty limitless sky in such a way as he had never seen it before! He has had a fleeting glimpse of the witness-consciousness—a fleeting glimpse, but *redemptive*. Needless to say, such a recognition of the 'Witness', the साक्षिन्, lies at the heart of the Upaniṣadic experience.

All told, here is a truly brilliant accomplishment, whatever way one looks at it. And to think of the entire series of volumes which are in the plan, one marvels at the sustained mental energy that must go into this enterprise of monumental proportions. Whether it will carry conviction to the 'Faustian man' who is ostensibly being addressed may be a matter of doubt—inured as he is to the frame of mind that the relentless march of modern civilization engenders. However, the ordinary man native to the Vedāntic soil, too, has good reason to be grateful for this largesse.

The main thrust of this enterprise will strike one as a triumph of the creative hermeneutical imagination. In effect it offers us an enormously enriched understanding of the *Mundaka* and the *Māndūkya Upaniṣads* and, what's more, a fresh outlook on Indological studies generally.

50, Yoginagar Society, Fatehganj, Baroda-2 V.Y. Kantak

Charles Prebish: *Buddhist Monastic Disciplines: The Sanskrit Prātimokśa Sūtras of the Mahāsaṇghikas and Mūlasarvastivādins*

The main title of the book is a bit misleading. It is not a treatise on Buddhist monastic discipline in general. Rather, it mainly presents the (first English) translations of Sanskrit Prātimokṣa texts of the Mahāsāṇghika and Mūlasarvāstivāda schools. These translations are prefaced by a *brief* introductory chapter on 'The Rise of Buddhist Monasticism'. This chapter discusses, among other things, the history of the rise of Buddhist monasticism, description of the basic texts of Vinaya literature, and ritualization-process of the Prātimokṣa formulary.

A survey of the etymological meanings of 'Prātimokṣa', suggested by scholars like Rhys Davids, Oldenberg and Sukumar Dutt, prompts Prebish to surmise that perhaps, at the beginning, the Prātimokṣa had nothing to do with offences committed by monks and their confessions. His search for the original form of the Prātimokṣa leads Prebish to verses

26, 27 and 28 of the Dīgha Nikāya (which are quoted at the end of the Mahāsāṇghika and Mūlasarvāstivāda Prātimokṣa Sūtras.) A scrutiny of these verses, which Prebish claims to be the earliest form of the Prātimokṣa, reveals that before being ritualized, the Prātimokṣa formulary had hardly anything to do with the listing of the monastic offences, their confessions and eventual expiation. Rather, they were, Prebish claims, concerned mainly with the declaration of faith in Buddhist principles, on the part of the monks, at Poṣadha ceremonies. At the end of the ritualization —process, the name, 'Prātimokṣa', came to be foisted on a code of prohibitions for a monk (incorporating also the formal declaration of faith).

The author claims that a careful study of the introduction (nidāna) to the recital of Prātimokṣa (by a competent elderly monk) discloses that the Prātimokṣa Sūtra is a formal and artificial rendering of a certain ritualistic formulary.

Although the elderly monk, reciting the introduction, praises the virtue of confession and emphasizes conscious lying to be a hindrance to the monk's progress, he also makes it clear that the first duty of the assembly of monks is to declare complete purity, *before* the start of the Prātimokṣa recitation (p. 25). The offences likely to be confessed during the actual recitation would, then, be such as those remembered at that time. However, such a confession is likely to be an exception. The introduction ends with stating that there is absolution for a monk who confesses a previously unconfessed offence (p. 25). Obviously, then, the possibility of an offence, for which a confession would not suffice for an absolution (such as a pārājika offence), is not seriously entertained at all (p. 25). It transpires that the recitation of the Prātimokṣa at a poṣadha ceremony is more a matter of formality. The occasion for the recital does not provide any scope either for the confession of offences necessitating immediate expulsion, or for the actual rectification of any offence committed.

However, the ritualization of the Prātimokṣa, Prebish, remarks, 'reveals not that ethics and morality were overlooked, but rather they continued—recast in the formalistic mold...' (p. 27) The Prātimokṣa, in his opinion, is not just a a monastic bond holding the saṇgha together as suggested by Sukumar Dutt (p. 20), but, rather, 'the common ground on which the internally enforced ethical life is manifested externally in the community' (p. 27).

The book presents the author's *carefully* done translations of the Prātimokṣa texts of the Mahāsāṇghika and Mūlasarvāstivāda sects on facing pages to facilitate comparison. Those interested in the study of the monastic practices of early Buddhism will find the book a helpful guide to their scholarly enterprise.

In spite of the obvious merits of the book, there are certain shortcomings worth mentioning:

The author does not offer any helpful suggestions in rendering the accurate meanings of many a technical term. Faced with the pressure to explain these terms, he would, at the most, simply quote the meanings suggested by others.

(i) Let us start with the technical terms of the eight categories of offences listed in the Prātimokṣa. The eight categories are: (i) pārājika dharmas, (ii) saṃghāvaśeṣa dharmas (iii) aniyata dharmas, (iv) niḥsargika—pāyantika dharmas (v) pāyantika dharmas, (vi) pratideśanīya dharmas, (vii) saikṣa dharmas and (viii) adikaraṇa—śamatha dharmas.

The author gives a brief account of the kinds of offences, penalties for committing them, and methods for resolving the charges of offences, listed under these eight categories of dharmas. However, he does not make any serious attempt to elucidate the meanings of the eight technical terms mentioned above. It may be a fact that an attempt to provide etymological meanings of the technical terms will not ultimately be of much help. Yet, it is possible that a search for the senses, which people employing these terms actually have in mind, may prove to be helpful. In any case, there must have been reasons for listing several items under one category, rather than another. And a search for such reasons would have given us some clues regarding the meanings of the technical vocabulary. It is a pity that the author does not give his independent opinion as to how one should interpret these terms.

(2) In stating the tenth niḥsargika-pāyantika dharma the author makes use of technical Sanskrit terms like 'vaiyyāvṛtyakaras' and 'ārāmikas' (p. 66), without making the slightest attempt to explain their meanings. He merely refers to explanations of them, offered by F. Edgerton in his *Buddhist Hybrid Sanskrit Grammar and Dictionary*.

Sometimes the author translates certain rules literally without trying to solve the apparent inconsistency involved in stating the rules. Consider for example, rules 50 and 52 belonging to the pāyantika dharmas.

Rule 50 states: 'Half-monthly bathing was declared by the Blessed One, Except at the right time, that—is a pācattika' (p. 54). Half-monthly bathing would hardly suffice for a monk during the summer, and yet this bathing rule is meant for the 'hot season' (p. 84).

Rule 52 states: 'Whatever monk should give, with his own hand, hard food or soft food to a male ascetic, female ascetic, male wanderer, or female wanderer, that is a pācattika.' (p. 84).

It is not clear why giving food to an ascetic and a wanderer would constitute an offence. Without additional explanatory comments, the apparent incompatibility involved in the account of the rules thus remains unsolved.

The author has a tendency to translate portions of texts without attempting to bring out the proper import of these translations. This is most noticeable in his translation of adhikaraṇa-śamatha dharmas (pp. 106–09). The seven adhikaraṇa—śamatha dharmas are not really monastic

rules. They represent a system by which offences may be resolved. However, the author does not really try to explain the methods recommended for resolution of these offences.

He translates the first method as: 'To a case worthy of resolution in the presence of, we shall impart settlement in the presence of.' (p. 107). In the short introduction to this category of adhikaraṇa-śamatha dharma, presented in chapter I of the book, the author explains that this method of resolution refers to saṃmukhavinaya. He adds that the 'Śamathakkhandhaka' of the Pāli Vinaya explains this by the presence is necessary for the resolution of the offence. Is it a special sort of individual, and a special sort of Saṃgha? What does the phrase 'in the presence of the Dharma and the Vinaya' actually mean here?

The second method of resolution is translated as: 'To a case worthy of settlement based on recollection, we shall impart settlement based on recollection of' (p. 197). The reader is left in the dark as to what sort of recollection is being talked about, and what part does such a recollection play in any settlement.

The sixth method is translated as: 'To a case worthy of settlement which covers over, as with grass, we shall impart settlement which covers over, as with grass' (p. 109). The reader is left in the dark, again, as to what a 'case worthy of settlement which covers over, as with grass', really refers to.

The book ends with a helpful concordance table comparing the two texts translated with other Prātimokṣa texts preserved in Indic languages and a carefully prepared bibliography.

Department of Philosophy and Religion RITA GUPTA
Visva-Bharati University

SURESH SHARMA: *Tribal Identity and the Modern World*, United Nations University Press/Sage Publications, 1994, pp. 215.

This book introduces us to a style of research in tribal cultures of India. Anthropologists, the privileged interpreters of the tribal world, have argued on issues concerning the 'field view' and the 'book view' of a culture. Many of them claim that the 'book view' presents dogmatic aphorisms rather than the empirical reality. One can say in defence that the empirical scientism of the 'field view' is banal in its utter concreteness. Blending the two views may help integrate the entirety of historical experience and thus a holistic view of culture. Happily, this book embodies both the views.

For the people already familiar with the tribal situation in India, this book will be of interest. In his Preface the author raises a number of

relevant questions: 'Does the survival and resilience of tribal cohesion signify the continuing power of an archaic social substance? Should the conceptual ground upon which the prevalent categorization and understanding of the nature of the tribal identity is predicated be re-examined? What constitutes the distinctive substance in the historical texture of relationships and sensibilities, in relation to tribal identity in India? What is the nature of reassertion of tribal identity in the modern context? And what kind of possibilities does this reassertion make available for a serious reconsideration of the problem of ecology and human survival, as also the nature and role of the modern state and modern development?' (p. 10) Answers to these questions are admittedly not easy. There are theoretical difficulties which the author outlines in the first part of the book. The second part examines the implications of the modern onslaught on the pre-colonial equations between social cohesion and political authority.

With carefully delimited scope for enquiry, the book presents a bold historical narrative of the tribes of central India. On the negative side, the socially informed data on tribes are thin. Anthropologists will search in vain a theoretically informed narrative of the evolutionary premise that defines 'primitive tribe' a pre-cultural humanity. There is a connection between history and evolution of tribal culture. The book seemingly underestimates the challenges of evolutionary theory. Here evolutionism has been dealt with perfunctorily and also disconnected from the trajectory of colonialism. In complex ways, evolutionism and colonialism are related. There is a Third Eye, the power of Christianity, which is responsible for the near cultural genocide of the tribespeople in all parts of the world. Social Darwinism and Christianity have been used, in one form or another, to justify the colonial expansion of European civilization. These crucial factors in the tribal identity problem the author does not address.

The case study of shifting cultivation (pp. 139–66), which supports the author's main thesis, is a valuable and useful contribution to anthropology and ecology. The 1930s controversy between Verrier Elwin and A.V. Thakkar on the tribal practice of shifting cultivation (p. 143) has been aptly mentioned, but not fully treated. Along side this controversy is the even starker evidence of the colonial policy of keeping the tribes at bay. Elwin, a Christian missionary, pleaded that the tribes must be kept in isolation. Thakkar Bapa, a nationalist, argued that the tribes are harmoniously assimilated into the rest of Indian population and they must be allowed to remain so. Elwin was glorified as a 'tribal hero'. Archival evidence demonstrates that while he showed familial love for the tribes and respect for the freedom fighters, he was tied in best with the British Raj.

Following Will's explanation (p. 180), the author wants us to believe that in pre-colonial India there was a sharp demarcations between 'tribal' and 'Hindu' and that their antagonistic identities lead to incessant

combat. What echoes in this explanation is the British policy of distancing the tribes from the rest of the Indian population. The introduction of anthropology, the western colonial science of man, as a discipline in Indian Universities and the decennial census of Indian population by caste and tribe are glaring examples. As the British record shows, the famous Risley, the hero of Indian anthropology, had recommended a tripartite division of India: Hindu, Muslim, and Tribal Nations. Ironically the post-Independence India was left with no option but to continue the British policy of scheduling the tribal population and restricting the tribal areas in the name of protection.

The author's reflections on the colonial discourse show that we the Indians have inherited a colonial mind from the British Raj. Elaboration of this view finds in the following statement: 'Through the long years of struggle against colonial rule Gandhi never ceased to speak of "Tribals, Harijans and Women" as the cardinal categories of suffering and injustice. But even that voice, incomparable in its effectiveness when it came to waging battle against colonial rule, seemed distant and unreal' (p. 75).

Underlying the 'Big Question' in tribal identity and the modern world is a general concern for the future of tribal communities. The modern world recognizes the tribes as having rights to life and provides an assumption for their gifted place in the universe. But its ceaseless chatter seems to hide a deep hatred, the horrific side effect of social Darwinistic thinking. Modernity is not a time concept. It refers to an ideology that encompasses the total human life and experience, a worldview diametrically opposed to what the traditional cultures have envisioned. Traditionally, tribal society is a complete world, a social universe, with its own character. Cosmologically it makes a 'perfect form' defined in terms of a self-generating, self-organizing, and self-perpetuating critical mass coherent with the internal and external environments. Modernity considers tribal society as a fringe of leaves—imperfect, primitive, unchanging, undeveloped. 'Modern man' appears in the edifice of Darwinian evolution. For him it is ethically right to destroy the primitive tribal cultural structures and to 'develop' the 'undeveloped'. Hence, the question of tribal identity today, even for the 'modernized' tribes, is irrelevant. Tribes may survive as a nation, not a culture. A frightening thought indeed that a 'perfect natural form' of human society is lost, only a memory of destruction remains.

This book is in many respects an interesting work. It contributes to a general understanding of how historical processes have resulted in the unmaking of a culture and constitutes an important contribution to the ethnohistory of the tribespeople of India. It is highly recommended for anthropologists, political sociologists, and historians working on tribal polity.

BAIDYANATH SARASWATI

N.K. DEVARAJA (ed.), *Philosophy of Religion*, Indian Institute of Advanced Study (IIAS), Shimla, in association with Indus Publishing Co., New Delhi, 1989, pp. 107, Rs. 100.00

The book under review is an anthology of six divergent articles concerned with vital currently discussed problems of 'Philosophy of Religion': self, self-identity, person, self-consciousness, hemeneutics, communal harmony, spirituality and the like. The contributors of the papers have tried to throw a fresh light on these perennial philosophical problems. These articles will provide new materials for the readers to reconsider them. All human thinking is action-oriented. If our theoretical articulation is not workable in practice that is of no use either. Therefore theory and practice are to be reconciled.

It is unusual for a reviewer of critical bent of mind to put in written words thoughts about stalwart scholars who are deemed to be established original thinkers. The reason for doing this uncomfortable exercise is primarily due to personal regards to the editor (Professor Daya Krishna). I will try to analyse the papers one by one with a view to exploring their philosophical significance. The title of the book is apt because the authors have dealt with philosophical issues though with religious fervour. Your reviewer does not find in the articles chiselled philosophical analysis of the problems. However there are some articles which deserve our appreciation for their indepth analysis. This will become clear when I take up the papers one by one.

Professor N.K. Devaraja, the editor of the book, discusses metaphysics of human life in the 'Introduction' (pp. 11–15). He starts with a sweeping statement that: 'Man's religio-metaphysical quest, it seems, began soon after the dawn of *civilization, expressing* itself in questioning about the nature and destiny of the self' (emphasis added) (p. 11). Why should we start with the dawn of the civilization? Even earlier the quest for unravelling the mystery of the self and human life had been carried out by the creative minds. So this creative probe into innermost recesses of human nature had been in vogue since the creation of the *homo sapiens* on the earth. Professor Devaraja has misconstrued that: 'The Upaniṣadic sages conceived the self as a 'substantive entity' or soul, that survived the death of the body and reaped the rewards and punishments for the good and evil deeds respectively done by it in its earthly existence in the embodied state' (p. 11). Here our philosophical problem is to correlate soul with the fluxional nature of the psychic life associated with it. Often it has been seen that religious traditions believe that human soul in essence is changeless, unborn and eternal. As a scholar of philosophy it is duty of the author to reason out the issues involved there in. Instead of analyzing the problem our author meaninglessly compares it with Hegelian view that: 'Hegel substituted the category of the "subject" for that of substantive soul and ruled that philosophy occupies itself with the description or

delineation of the life of the spirit conceived as a developing subject' (p. 12). Such a mature thinker should have tackled the problem of the identity of the self of its own from the Vedāntic standpoint rather than side tracking the main issue under discussion. Hegelian subject and the Upaniṣadic *ātman* are poles apart and thus incomparable.

Again we will find Professor Devaraja's tailed article: 'The Humanistic Approach to Hindu Religio-Philosophic Thought' (pp. 101–07). Your reviewer fails to see eye to eye with the title and the contents of the paper and his approach of the Upaniṣadic thought as humanistic. Basically humanism was an atheistic viewpoint of human values without divinity and transcendent Being. I do not think Jean-Paul Sartre, Marx, T.S. Eliot, *et al* could be branded as humanists in this sense of the term because the transcendent Deity/Being is missing in their writings. So Indian or Upaniṣadic spiritualistic thought, on the contrary, can be better termed humanitarianism rather than humanism. In humanism the central metaphysical nucleus is man whereas in humanitarianism there may be divinity in the centre and human values may reel around it. Our author without appreciating the transcendental philosophical significance of the Vedāntic thought tries to impose western atheistic approach of humanism on human life.

Furthermore, Professor Devaraja fallaciously argues that the concept of *jīvan-mukta* in Advaita Vedānta, Sāṁkhya-Yoga and *Gitā* is clearly in line with a humanistic philosophy of life. *Jīvan-mukta* is living *in* this world but he is not *of* this world. He inwardly transcends all attachments of this-worldliness. Here is Eternity yoked in temorality or history. How can we call it humanistic viewpoint of life philosophy? This argument will not work that: 'liberation can be worked out and attained only by human beings, neither by lower animals nor yet by gods' (p. 106). And if we go deeper into the Hindu doctrine of *mukti/mokṣa*, it will be revealed by implication that human self (*ātman*) is also emancipated from personal Deity (*Īshvara*). Personal Deity is a hypostesized intellectual construct symbolizing the transcendent Being. He disappears as soon as *mukti* is attained. In Indian Tradition Personal Deity does not have an ontological reality as we find in the Semitic Traditions.

It is a fantastic interpretation that: 'Upaniṣadic Vedāntic outlook on life is frankly individualistic' (p. 107). The role of the *saṁyāsins* in the society is that of an altruist and he is out for the emancipation of the all and sundry. He is a self-realized man living in the society with disinterested involvement for the emancipation of the living beings. So humanitarianism in the Indian tradition is qualitatively distinct and has a wider connotation in comparison with the European atheistic movement of humanism.

First Debabrata Sinha's paper, 'Towards a Philosophical Anthropology From A Vedantic Perspective: A Hermeneutic Exploration' (pp. 17–42), is a nicely articulated one. It starts with Max Scheler's significant quotation from his work, 'The Place of Man in The Universe', indicating man's

capacity to surpass himself the world. As compared to lower species man has an inherent force to go beyond himself. Thus this faculty of self-transcendence is a distinctive urge of theoretical articulation through words or language to resolve practical problems. This is the essence of Indian thought of man's capacity for self-transcendence. It echoes Max Scheler's 'contemporary philosophical anthropology, repeats the western account on the human being as a person's creative drive' (Karl Jaspers).

This is the main reason that man has been privileged to attain *mokṣa*. This serves as a basis for anthropological inquiries. It is said that no serious conceptual attempt has been made to study man as a unity of body and soul. But the Bhakti cult of India in the medieval period, initiated by the Tamil Siddhas (ninth century), the body has been treated as a boat for realization of *mokṣa*. Human person is treated as a microcosm (*piṇḍa*) reflecting the macrocosm (*Brahmaṇḍa*). It is an 'attempt to place the human being in the hierarchy of living beings of the physico-biological order, to relate human reality to the order of realities and events in the larger world' (p. 19). Culmination of the development of consciousness in cosmic consciousness yoked in historicity is human life's development (*mukti*). On the other hand, the Vedāntic perspective treats human body as *māyā* (illusory). It is an abstract perspective transcended from the temporal world (*jagat*).

The second section (Methodological Prolegomena to a Hermeneutic Self-understanding Across Philosophical Cultures) (p. 20) deals with subjective understanding of human self. In this article Sinha correctly says hermeneutical understanding is a 'way of interpretive understanding its originary insights. . . .' (p. 20). The process of accumulating historical consciousness adds new dimensions. Indian scholastic tradition of Sanskrit *paṇḍits* always look from insider's view, i.e., 'repetition rather than an historical understanding' (p. 21). It is a prerequisite of hermeneutics of internalization of interpersonal subjectivity. Development of consciousness projects itself in the external world as well. Hence pristine insights are to be understood alongwith accredited historical consciousness developed through the tradition.

In Greek thought man is mortal and temporal, and has no link with Eternity. He cannot transcend this-worldliness. Nietzsche got this idea of tragic view of human life from the Greeks (see, *The Birth of Tragedy*). On the other hand, Indian concept of self-transcendence is a redeeming feature of optimistic and dynamic view of human life. This is the spirit of human life in Indian philosophy.

Next section (III) is about 'Contemporary Philosophical Anthropology' leading to, '. . . search for the essence of the human being'. Our author is fully justified to assert that, 'the fundamental phenomenon of human existence and. . . the human condition seek to emphasize that human existence is never an entity in itself' (p. 27). It is subjective and fluxional in nature. For this reason Martin Heidegger places his Being in time. But

basic problem is: How to reconcile 'human subjectivity vis-a-vis the transcendental ego of Husserlean phenomenology' (p. 28). This is a perennial problem of philosophy in which man is a bridge between transcendental Self (Eternity) and this-worldly self. Section IV is an abortive attempt to encompass Eternity and temporality within the Vedāntic model of *cit*-centricity. In the fifth section we find break through towards a hermeneutic anthropology in the Vedāntic *ātmanic* perspective. Here *māyā*, *avidyā* and *ajñāna* are negative categories. Their hermeneutical significance is taken up. Our author has correctly pointed out that, '*avidyā* is not attributed to any extraneous "transcendent" source other than the immanental life of consciousness' (p. 38). It requires new hermeneutical methodological framework of conceptual scheme. From this standpoint new thematic structure of 'anoetic principle of *avidyā* is projected through an eidetic description-cum-analysis of consciousness' (p. 38). Pure consciousness (*ātman*) is 'noetically-eidetically transparent' (p. 39). From epistemic and praxeological standpoints *śakti* as a dynamic principle is not only conative manifestation of *Brahman* (*Śiva*). It has twofold dimensions—cognition and volition. This pure cognitive-will is the real *vis a tergo* of human dynamism. This is how consciousness through volition is to be embedded in temporality. I enjoyed reading this paper. Let me hope the readers will also have the same experience when they go through it with meditative concentration.

Next we have Jim Stone's paper 'Parfit and The Buddha: Why There are no People' (pp. 43–58). It draws our attention on the ontological status of person. It raises a question: Is person only body and mind? Derek Parfit is of the view that: 'persons are not something extra' (p. 43). Then the question arises: What is the identity of the person? Parfit takes the position of reductionism and explains it with phenomenological experience. The author argues from the point of view of elimination which is a form of reasonable choice. Again in the Part III he shows that 'Reductionism is incoherent: either persons are extra or there aren't any. There is no middle ground' (p. 44).

In Part I he takes up the problem of personal identity and makes a distinction between heap and structure. For him 'people are psychological structures, not heaps' (p. 46). His spectrum argument concludes that 'personal identity is inderminate supports Eliminatism...' (p. 47). There is no such an eternal substratum or substance which can be called personal identity symbolizing self ('I').

Part II raises a vital philosophical question: Is identity relationship of the person synchronic or diachronic? The underlying ontic identity is synchronic or similarity. There is no connectedness of psychological events. Hence reductionism in all forms is false. Continuity is to be identical and not duplicate. In the fluxional nature of the personal identical duplication is a must in forms and the essence remains identical.

In the concluding Part (IV) our author has taken a stand that: 'There is no handy middle ground between Realism and Eliminativism' (p. 55). In the Humean sense it could be called 'a metaphysical fiction of substance.' As a matter of fact the created person is a *natura naturata*. The personality of the ego has unity of persisting series of diverse stages. These stages can either be vertical or horizontal or both at a time. This is inward journey of the self to realize the eternal through externalization in history. It also poses a Platonic and Upaniṣadic philosophical question of one and many. How is person one and many? This paper reopens the perennial debate of Hiraclitus, Buddha and David Hume to find personality of the self. Is the person beyond brain, body and psycho-physical entity? On the whole the paper is thought-provoking and is nicely argued.

Professor S.G. Tulpule's paper, 'The Genesis and Nature of the Divine' (pp. 59–71), is about the divine Name. The divine Name as a process of constant identification is called '*nāmasimrana*'. The divine Name also symbolizes the ontological '*isness*' of non-dual Reality. It is central in all religious traditions. The paper is an exegetical exploration with no philosophical and critical analysis of the concept of Name which could probe into its significance from mystical meaning of the term. The author is concerned more with history of Name rather than vertical hermeneutical development of the self and understanding of concept of Name. There is no evidence in the paper that may explain its philosophical implications in quest of reality and mysteries of human life. Our philosophical questions are: What is spiritual life? What is the aim of human life? How can human ego be transformed into a realized self? How does *nāmasimrana* function on man's consciousness? Is the sense of human life in Indian tradition tragic or optimistic? This type of analytical discussion is lacking in the paper. There is more of religiosity rather than a critical attempt on the topic.

Hugh van Skyhawk's paper, 'The Heart of Religion: A Sufi's Thoughts on the Relations Between Religious Communities', sees in the unity of God and moral imperative for the unity of man (pp. 73–88). Both exoteric and esoteric, or inward or outward approaches in Hinduism and Islam self-surrender of the individual will to the Divine Will lay emphasis on the eradication of 'the *ahaṅkāra* is the cause of the consciousness of individual existence' (p. 79). This paper will be of great interest to the devotional exegetes as no serious philosophical issues have been touched in it.

The last paper, 'Religion, Virtue and Spirituality' (pp. 89–99) by M.M. Agrawal expounds final goal of freedom and pathway to the genuine religions. In this scheme of the religious traditions we find that 'virtue, spirituality and religion are linked together' (p. 90). The goal of all religions is freedom of the self. Jains have ascetic or self-mortific view of life and Buddhists advocate middle-path. The goal of moral life is wider than the goal of a religion. This way leads to a 'somewhat paradoxical

situation' (p. 91) between egoism and altruism. This is an apparent paradox, but it is resolved when one's egoity (*ahaṅkāra*) is dissolved into interests for others. This idea has been dealt with in length by Indian religions. Our author has shown utter ignorance about the core problems of human existential conditions. It was duty of the editor to point out in it in the form of footnotes. This glaring lapse on the part of the author minimizes significance of the paper.

Spiritual and religious aspects of faith are different. To be religious is not necessarily ethical and spiritual. Priest or devout devotee is religious outwardly but inwardly he may lack spirituality and morality in deeds. So morality taken in its true spirit culminates in spirituality, i.e., freedom of the self which is its intrinsic ontic value. Hence religious aspect of faith is ritualistic and formalistic, and spiritualistic one is genuine and intrinsically good. The author has not clearly shown sharp distinction in the concluding remarks on p. 91. He leaves the problem as a problem without making a least attempt to resolve or dissolve it.

In the section 'The Structure of Religion' (p. 93) faith, worship and obedience are taken as key-terms. This term is a prerequisite for personal deity. The author has correctly pointed out: '. . . religion requires is not necessarily God but a belief in a transcendental reality which is considered as sacred and which somehow contains in itself the clue to the mysteries of existence' (p. 93). When the mystery is unravelled, the hypothesized personal Deity disappears. This is how I think about the essence of all genuine religions. In Buddhism and Jainism there is no such a personal Deity. Hence in these traditions emphasis is on meditation (*samādhi*) and concentration (*dhyāna*) and austerity (*tapas*). Sikhism chooses a via media between the theistic and non-theistic spiritual traditions.

In Hinduism when there is no intermediary charismatic personality God incarnates Himself and reveals in words and deeds the true values and ideals for the mankind. Sometimes a sacred literature is deliberately turned into a revealed religion. The author did not discuss the nature of Semitic revelation (*waḥy*) and the Indian concept of revelation (*śruti*). Both have different natures, origins and functions. The author has uncritically assumed prophethood and *avtāra* as equivalent and identical concepts. To my mind it is a category mistake. Both the concepts have different roles to perform in their respective religious structures. Here the author's observation is of a layman rather than of a critical and analytical thinker.

Our author points out that: '. . . I am not denying that a religion can exist without any conception of God or a transcendent reality' (p. 94). I think God or transcendent reality is a mental or rationally articulated construct by man's creativity to grapple with existential human conditions. When the problems are dissolved in theory and praxis, their spiritual constructs are done away with. In the Semitic traditions personal Deity as a transcendent personal Being always remains the 'Holy Other', but in Indian traditions the self (*jīva-ātman*) and the personal Deity (*Īśvara*) are

dissolved into the primordial Formless Unity (*Brahman*). Our author should have drawn out such a vital distinction between both the traditions.

Lastly the author concludes that: 'religion is the means, spirituality is the end' (p. 96). What is that spirituality? Our answer is: It is the Holy or the Sacred. This transcendental reality is the ground of the phenomenal reality. The spirituality of human life consists in somehow going beyond the phenomenal world. The author has correctly pointed out that spirituality is not a transcendental metaphysical reality but a state of emancipated consciousness from suffering and realization of self's 'ontological status of this freedom' (p. 97). Spiritual freedom culminates by virtue of selflessness. 'In virtue there is a transcendence of the self' (p. 97). It is, therefore, one's interests and others' interests are reconciled in true love (*prem*).

On the whole, the book contains wisdom of the renowned contributors in the field of Philosophy of Religion. The present anthology of current topics will definitely draw attention of the serious renders to develop their deeper interest in the religious issues.

NIRBHAI SINGH

Obituary Notes

Surendra Sheodas Barlingay
20 July 1919–19 December 1997

The death of Dr Surendra Sheodas Barlingay, Emeritus Professor of the University of Pune, on 19 December 1997, brought to a close a distinguished career of varied interests and removed from our midst a very lively personality whose impress has been left on several spheres of activity in philosophy, education, politics, literature and culture. He died at the age of seventy-eight leaving behind a formidable corpus of philosophical writings, a large institutional network which he had set up in the field of education and above all, innumerable colleagues and pupils from almost every part of this vast country.

He was born in a leading and noble family from Nagpur. His father, an advocate of high repute, was also a renowned scholar of Sanskrit by virtue of his being a *Mahamahopadhyaya.* Professor Barlingay's elder brother, Dr W.S. Barlingay was a Cambridge-trained philosopher and a student of G.E. Moore and C.D. Broad. Instead of pursuing their professional careers both Professor Barlingay and his brother responded to a call given by Mahatma Gandhi and took active part in the freedom movement. In the charged atmosphere of the 1942 movement and under the influence of his father and brother, Professor Barlingay soon became an activist in social and political spheres. The period between 1942 and 1945 was the most active phase of his life and was of a vital significance for him. During this period he came in contact with such national leaders as Achyut Patwardhan, Dr Ram Manohar Lohia, Jaiprakash Narayan and Yusuf Meherali. They were not merely agitators but thinkers in their own right. Discussions with these scholars, as Barlingay has acknowledged, helped him to understand the theoretical and practical sides of our social problems.

Amidst the thick of political activities, Barlingay's career in philosophy began with his fellowship at the erstwhile Indian Institute of Philosophy at Amalner. There he came in contact with Dr Ras Vihari Das, Professor G.R. Malkani, Professor D.Y. Deshpande and Professor Chandrodaya Bhattacharya. To some extent, the fellowship period at Amalner served as foundation for his philosophical studies. His professional life as a teacher of philosophy dates back to a lecturership at Raipur during 1945 and 1948. But even during this short-term assignment he could start a weekly called *Inquilab* and run it successfully for some years. But the turning point in his academic career came when he met Swami Ramanand Tirth, a great freedom fighter who was then the President of the

Hyderabad State Congress. At his behest, Barlingay left Raipur and went to Hyderabad to establish a network of educational institutions spread over the former state of Hyderabad. He served as founder Principal of these institutions. The centre of his activities during this period was Nanded. Amidst these activities of institution building and participation in the Hyderabad freedom movement Barlingay wrote his doctoral dissertation on the 'Concept of Change' for which Nagpur University awarded him PhD in 1956. In the same year he left Nanded and went to England for his post-doctoral studies under Gilbert Ryle. Barlingay's acquaintance with P.T. Geach, Anscombe, Dummett, Austin, Cohen, Neil and Mrs Neil, Hao Wang and others, changed his style of philosophizing. During the same period, a number of logicians of repute were in England as visiting professors whose lectures proved to be extremely beneficial for Barlingay to develop his own understanding of diverse branches of logic such as predicate logic, axiomatic theory, theory of sets, modal and many-valued logic and intuitionist logic. His convivium with Brower and Stall centred round the two principal strands in logic, namely, formalism and intuitionism. It is with this equipment in various branches of western logic alongwith his training in Indian logic especially in Nyāya and Navya-Nyāya which he had received from Shri. N.S. Rajgopal Tatachari and D.T. Tatachari in Tirupati where on his return from England he served in S.V. University as Reader, that he could write his first major work *A Modern Introduction to Indian Logic* while he was in Zagreb in 1962 as Professor of Indian Philosophy. In 1964 he joined Delhi University as Reader and Head. With the help of Professor Amartya Sen he introduced a course in Philosophy of Science, probably for the first time in this country. In 1968, Professor Barlingay went to the University of Western Australia as a visiting Professor and remained there till 1970. He joined Poona University in that very year and stayed till he retired in 1980. His presence in Poona University is marked by many achievements, two of which are well-known—the establishment of three journals—*Indian Philosophical Quarterly, Paramarsha* in Marathi and in Hindi, and also the creation of a post of a Nyāya Pandit. The community of philosophers in India will always remain grateful to him for establishing these journals at a time when there was hardly any journal for publishing scholarly articles.

Another significant event in his life after retirement from Poona University was his appointment as the Chairman of the State Board of Literature of Culture, Government of Maharashtra. During his seven years' chairmanship he introduced many new schemes especially for poets and writers from the weaker sections of our society. Many schemes which he has introduced triggered off controversies and invited criticism from the press. In retrospect, however, the merit of these schemes is slowly but definitely being recognized.

Since 1988 Professor Barlingay remained mainly in Pune and finalized

many of his research projects and other writings under the auspices of various fellowships from the Indian Council of Historical Research and Indian Council of Philosophical Research. He was remarkably productive and despite his ailing health he was active till the last day of his life. In 1995 he had organized a World Philosophers Meet in Pune, which was an event to be remembered. His last official assignment was as a volume editor for a volume on 'Indian Aesthetics' which he was preparing for the Project on History of Indian Science, Philosophy and Culture.

Professor Barlingay was among the few contemporary Indian philosophers who have earned almost every honour of excellence in philosophy. He will be remembered by his many friends not only for his institution-building, his intellectual gifts, his kindness and his personal courage, but also for his zest. To those who had the privilege of working with him, his outstanding qualities were his charm of manner, his initiation and ready reception of new ideas, and the great encouragement he gave on every occasion to younger men, particularly when difficulties arose. It is for the future generations of philosophers to critically assess his contributions to philosophy; but for his contemporaries, and members of his family, it is a personal loss of a genuine human being.

University of Pune SHARAD DESHPANDE

Professor Surendra S. Barlingay: philosopher, visionary and activist

In the passing away of Professor Surendra Barlingay, the world of philosophy has lost an eminent thinker and creative interpreter of Indian tradition and culture, the academia lost a dynamic organizer, and generations of students lost a teacher, guide and philosopher. Amiable and affectionate by nature, he was a maverick thinker who came to be known for his refreshingly original ideas and for the rich contribution that he made to philosophical thinking in the recent times. Barlingay combined in him the dynamism of a social activist, the vision of a institution-maker, and the vibrancy of an articulate thinker for whom philosophy was both a passion and a way of life. A robust sense of humour and a motherly compassion that he felt for everyone around him lent to his personality almost an irresistible charm. Like a collosus he dominated the Indian philosophical scene for nearly the last four decades. During this period, he taught, lectured, organized academic events, presided over research and academic bodies, founded colleges and research institutes, edited two important philosophical journals in India, and worked for setting up the Peace University at Alandi in Maharashtra, in

collaboration with MAEER's M.I.T, Pune. His philosophical writings will be a beacon of light to thinkers and scholars in the times to come.

Surendra Barlingay spent his early years in Nagpur. In 1928, when the Simon Commission visited India, as a nine-year-old boy he took off his cap and threw it into the bonfire of foreign goods. Soon, he started wearing khadi and was drawn into the vortex of the freedom movement for the country. In the early 40s, as a great upsurge of patriotic feeling ran through the length and breadth of the country he joined the students' movement and later became Secretary of All India Students' Conference (Nagpur session). Earlier, as a young boy of 15 he had been arrested by the reserved mounted police for distributing the Congress bulletin and was dragged to police custody like a dead dog. In 1942, he was in the thick of the Quit India movement and remained mostly underground till 1945. On one occasion he escaped immediately after being arrested by the police. A recipient of the Freedom Fighter's pension, he felt more proud of this token of recognition than for all the accolades and honours he had received for his contribution to philosophy.

Cast in the mould of a Renaissance man, Professor Barlingay had wide-ranging interests and delved deep into many areas of knowledge. He swept aside the narrow demarcation lines of academic specialization in his dogged search for the linkages and the underlying thread of unity among the various disciplines. No arm-chair thinker in the ivory tower, he drew enormously from his own experiences in life which he lived in its myriad fullness. He was as much at ease with the front-runner scientists as with, say, a group of grass-root political workers or social activists. He could befriend anyone with stunning ease.

Communicating with others was, for Barlingay, a way of attuning himself to the process of social dynamics. He laid great store by human communication as he believed it to be the antidote to a fractured world where each individual remains alienated from the others. In his writings on social and political philosophy, he drew a distinction between what he termed as the cosmocentric world and the anthropocentric world such that the latter was treated as 'a construct over' the former, the underlying idea being that it is possible to build a 'human world' only through human intervention. An ardent believer in the *Sarvodaya* model of Vinoba Bhave, he pointed to the great promise it held for leading us toward a world where 'man becomes aware of himself and becomes aware of others as equal to him'. His own vision of 'One World' was closely linked with the idea of community-living which he thought could be ushered in by the process of decentralization at different levels.

Complementary to such a conception of human society, Barlingay developed on the idea of human freedom which, on his view, should not be related with or to the individual or society in exclusion from each other. Instead, the concept of freedom should be understood in the

context of the 'individual-in-a-society' so that 'freedom' and 'duty' would be regarded as two sides of the same coin.

Professor Barlingay's deep convictions about life, world, and human action had much to do with the rough and tumble he went through in the formative years of his life. His own conception of philosophy was closely linked with the need for having a two-way process which involves doing action on the basis of reflections and doing philosophy in relation to such action. Philosophy, for him, 'would be theorization of what man does, for man is nothing but a gestalt of his actions—self-conscious actions and a critique of his actions'. Philosophy without action would have no significance just as human action without philosophic reflection would cease to have its creative orientation.

Viewed from such a perspective there should be no antagonism between a life of action and a life of reflection. Following along these lines, Barlingay succeeded in no small measure in pulling academic philosophy from out of the narrow confines of the class-room and library in order to focus on its wide ramifications in vairous spheres of life and activity. He often lamented the lack of awareness among academics and others as to the proper role of philosophy which, according to him, is that of setting the agenda for social reconstruction. Instead, philosophy is confused with theology, religion, and mysticism because of which it is often treated with contempt and scorn.

In his presidential address at the Indian Philosophical Congress (Madras session, 1979) he made a carping criticism of attempts to marginalize the role of philosophic pursuit: 'They do not understand that philosophy is the very life-blood of man, that it penetrates every action of man, that it cannot be separated from him'. Even the social ills of corruption, lack of concern for others, and immorality in public life can be tackled on a sound footing by encouraging an attitude of self-criticism and reflective self-consciousness. Barlingay's writings on these issues breathe his firm conviction that philosophy should not be divorced from life of action as the two play a complementary role to each other.

Deeply grounded as he was both in Indian and western traditions, Professor Barlingay wielded his logical tools unsparingly while revisiting many of the important philosophical issues and concepts. He believed that many of the problematic issues that one finds in Indian or western philosophical traditions owe their genesis to a confusion between the linguistic substantive and the physical substantive. According to him, a proper way to deal with such matters would be to understand the divide between what he termed as 'distinguishables' and 'separables'. The notion of separability is intelligible only in the context of things and physical existents. When a thing is divided into two or more parts each of these may exist separately as a physical entity. This would not be so if we are dealing with logical abstractions. What is distinguishable in

thought need not necessarily be divisible as, and into, physical substantives. Barlingay argued that when Bertrand Russell and others talked about 'sense-data' it was, as a linguistic substantive, only distinguishable and not separable from physical things and therefore, to look for sense-data as physical entities would be really to confuse between the two different categories.

In a similar vein, he held that Space and Time are not separable from things but are only distinguishable from them. The argument was extended to refute the Vaiśeṣika view that space and time are independent substances. The Nyāya position on *Dravya* and *Guṇa* also came under similar scrutiny. It was Barlingay's view that a proper appreciation of the logical divide between 'distinguishables' and 'separables' could be used with advantage for understanding many of the problems and paradoxes that have regularly ensnared philosophers in different times and climes.

With an abiding interest in the search for the roots of Indian tradition and culture, Barlingay worked toward developing an integrated vision which would combine the logico-epistemological, ethical and aesthetic aspects of Indian thought. His monumental work on logic, *A Modern Introduction to Indian Logic*, was the first step in this direction which drew wide critical acclaim from scholars and thinkers. The second in this series, *A Modern Introduction to Indian Ethics*, came out about a year ago. And before his death last month, he was still busy giving the final touches to his manuscript on Indian aesthetics which was to be the last of the trio in series. His other important books on philosophy include *Beliefs, Reasons, and Reflections; Poverty, Power, Progress; A Critical Survey of Western Philosophy (From Bacon to Kant)*; *Kala aur Saundarya* (in Hindi); *Tarka Rekha* (in Hindi); and, *Saundary Tatwa aur Kavya Siddhant* (in Hindi). He also published several novels and short stories in Marathi.

During his long career as a professional philosopher, Barlingay taught at several Indian and foreign universities and visited Europe, America and the Far East in that connection on many occasions. On his return from Oxford in 1959, he taught for some time at S.V. University, Tirupati, from where he moved to Delhi University in mid 60s, and finally joined Poona University in 1980. In between he held visiting assignments at University of Zagrev (Yugoslavia), and Western Australia University. He remained associated with Poona University till the end as Professor Emeritus.

Apart from all his academic achievements what stood out prominently was a strong sense of commitment that Barlingay felt towards certain cherished ideals. As a social visionary, he could see beyond his time. And the idea of 'One World' was perhaps at the centre of such a vision: 'It is not the concept of one physical world but the concept of one human world. This can be achieved only when man becomes aware of himself and becomes aware of others as equal to him'.

The last few years of his life were taken up by his pursuit of working out the theoretical structure of such a world. For the same reason, the idea

of setting up of the World Peace University at Alandi (Maharashtra) remained so dear to him and he worked ceaselessly for the achievement of this objective in association with a well-known educational institute of high profile: MAEER's M.I.T.

All those who had the opportunity to come close to him would naturally feel a void now. Barlingay was a source of inspiration to many. For them his following words carry the most important message: 'Man is now alienated from himself. He is lost. We may talk of values but even values have become external to us. We have forgotten that man is the source and the measure of all values. Let it dawn on us that we have to find out this lost man. Philosophy is a search for this man'.

Department of Philosophy, University of Delhi RANJAN K. GHOSH
(former Director, ICPR, New Delhi)

ANNOUNCEMENT

The *JICPR* proposes to discuss the following topics in the forthcoming issues of the *Journal* and would like to invite papers dealing with them from scholars interested in the areas mentioned:

1. Modernity and Post-modernism: A Critical Evaluation
2. Philosophy and the Social Sciences
3. Philosophy and Literature

The *JICPR* would also like to invite its readers to suggest topics that they would especially like to see discussed in the pages of the *JICPR*. Those interested may write to the Editor in this connection.

EDITOR

Journal of Indian Council of Philosophical Research

Editor: DAYA KRISHNA

Volume XV Number 3 May-August 1998

ANNOUNCEMENT

The *JICPR* is seriously thinking of forming a Network Group consisting of those of its readers who would like to receive the contents of its special features such as 'Focus', 'Agenda for Research' and 'Notes and Queries' before their publication so that they may become aware of them as soon as they are received and may respond to them in case they would like to do so.

The *JICPR* is at present published three times a year and thus it takes a long time for items under these sections to be published and brought to the attention of our readers. In order to avoid the delay, it is proposed that those who would like to be actively involved in the on-going discussions may write to us expressing their desire to become members of the *JICPR* Network Group so that they may be sent the material immediately as soon as it is received by us. Those interested may kindly write to the Editor.

EDITOR

ORDER FORM

Dear Sir,

Please enter my annual subscription for becoming a member of *JICPR*'s (*Journal of Indian Council of Philosophical Research*) Network Group. I enclose a cheque/DD for Rs. 50 in favour of ICPR.

Signature_____ Date_____

Name_____

Institution_____

Address_____

INDIAN COUNCIL OF PHILOSOPHICAL RESEARCH

Forthcoming Publications

Jñānagarbhā's Commentary on Just the Maitreya Chapter from the Samdhinirmocana-Sūtra: Study, Translation and Tibetan Text
John Powers

Ludwig Wittgenstein's Culture and Value, Second edition
translated into Hindi by *Ashok Vohra* (with text in English on facing pages)

Ludwig Wittgenstein's On Certainty
translated into Hindi by *Ashok Vohra* (with text in English on facing pages)

A Glossary of Technical Terms in the Commentaries of Śaṅkara Rāmānuja and Madhva on the Brahma-Sūtras Part II
K. Jayammal

Language, Testimony and Meaning
Sibajiban Bhattacharyya

Fundamentals of Logic
Arindam Singh and *Chinmoy Goswami*

ANNOUNCEMENT

The *Journal of Indian Council of Philosophical Research* plans to publish an issue dealing with implications of developments in modern logic for philosophy under the editorship of Prof. Anil Gupta and Dr Andre Chapuis, under the title "Circularity, Definition and Truth". Papers are invited for inclusion in this issue. Authors should send their manuscripts to the following address:

Dr Andre Chapuis
Department of Philosophy
Indiana University
Bloomington IN 47405
USA

The deadline for the submission of manuscripts is 1 July 1998. It is expected that the issue will appear some time in the year 1999.

Diacritical Marks

Vowels

आ ā
ई ī
ऊ ū
ए,ऐ ē } (long)
ओ ō } (N.B. long ē and ō are for the particular syllables in Dravidic languages.)
ऋ ṛ and not ri; (long ॠ, which rarely figures, may be rendered as ṝ)

Nasals

Anusvāra
(ं) ṁ and not ṃ

anunāsikas
ङ् ṅ
ञ् ñ
ण् ṇ (or ṇa as the case may be)

Hard aspirate

Visarga
(ः) ḥ

Consonants

Palatals
च ca and not cha
छ cha and not chha

Linguals
ट ṭa
ठ ṭha
ड ḍa
ढ ḍha and not ḷha

Sibilants
श śa
ष ṣa
स sa

Unclassified
ळ ḷa
क्ष kṣa and not ksha
ज्ञ jña and not djña
ऌ ḷṛ and not ḷri

General Examples
kṣamā and not *kshamā, jñāna* and not *djñāna, Kṛṣṇa* and not *Krishṇa, sucāru chatra* and not *suchāru chhatra* etc. etc., *gaḍha* and not *gaḷha* or *garha,* (except in Hindi)

Dravidic (conjuncts and *specific) characters*
ൽ ḷ
ഴ ḷ.
ൺ ṇ
ற ṟ.

Examples
Iḷaṅ-Gautaman, Cōḷa (and not Chola),

Muniṟpuvamangalam, Māran etc.

Miscellaneous

Where the second vowel in juxtaposition is clearly pronounced:
e.g. *jāṇaï* and not jāṇai
Seüṇa and not Seuṇa

Also, for English words showing similar or parallel situations:
e.g. Preëminence and not preeminence or pre-eminence
coöperation and not cooperation or co-operation

For the Simhalese, excepting where the words are in Sanskrit, the con-ventions of rendering Simhalese in Roman are to be followed:
e.g. *dāgaba* and not *dagaba*
veve or *véve* and not *vev*

Quotations from old Indian sources involving long passages, complete verses etc., should be rendered in Nāgarī script.
(The western writers, however, may render these in Roman script if they wish; these will be re-rendered in Nāgarī if necessary, by the editors.) Sanskrit quotations rendered in Roman are to be transliterated with *sandhi-viccheda* (disjoining), following the conventions of the *Epigraphia Indica*, but the signs for
laghu-guru of the syllables in a meter (when the citation is in verse) are not to be used.

Place Names

These are to be diacriticised, excepting the anglicised modern:
Examples: Mathurā, Kauśāmbī, Valabhī, Kāñcī, Uraiyūr, Tiḷevalli etc., but Allahabad (not Allāhābād), Calcutta (not Calcaṭṭā), Madras (and not Madrāsa).

Annotations

There will not be footnotes; but annotations (or notes and references), serially arranged, will appear *en masse* at the end of the text in each article.

References to published works

Those pertaining to articles, books etc., appearing in the main body of the text, or annotations, or otherwise:
Title of Book, Author's name (beginning with his initials) title, edition (if any) used, the name of the series (if it appears within it); next the place of publication along with year of publication, but without a comma in between; finally the page (or pages) from where the citation is taken or to which a reference is made.