ANDRÉ FUHRMANN

REFLECTIVE MODALITIES AND THEORY CHANGE*

1. INTRODUCTION

In 'Belief Revision and the Ramsey Test' (Gärdenfors, 1986) Peter Gärdenfors has shown that a prima facie plausible acceptance condition for subjunctive conditionals is incompatible with a certain apparently equally plausible requirement on how a rational agent should change his beliefs in the light of new information. The above mentioned acceptance condition for conditionals is the so-called Ramsey Test.

> *Ramsey Test for conditionals* (RT)
> A conditional $A > B$ is accepted in a state of belief K if and only if B is accepted in a state K' which results from K by minimally changing K so as to consistently include A.

The requirement on rational belief change which is incompatible with the Ramsey Test is the

> *Preservation Criterion* (P)
> If a proposition A is consistent with a state of belief K (i.e., $\sim A$ is not in K), then nothing has to be removed from K in order to obtain from K a state K' which differs from K only in as much as is needed to consistently include A.

Gärdenfors has suggested that the conflict between (RT) and (P) should be resolved in favour of (P).

In 'Iteration of Conditionals and the Ramsey Test' (Levi 1988), Isaac Levi argues that (P) does not only conflict with the way conditionals are generated into states of belief by way of (RT), but also with the presence in belief states of appraisals of serious possibility generated by an analogue of the (RT). One may think that this is a good reason for calling (P) into question after all: (P) appears as a disturbing factor in at least two different contexts. But Levi emphasizes what, in his view, conditionals and appraisals of serious possibility have in common.

Synthese **81**: 115–134, 1989.

> On my view, counterfactuals have truth values only insofar as they are construed as descriptions of the agent's conditional evaluations with respect to serious possibility. The appraisals themselves lack truth values and, as a consequence, so do the conditionals construed as expressions of such appraisals.[1]

Thus, at heart, a conditional is just an appraisal of serious possibility with respect to a state of belief revised to include the antecedent of the conditional in question. A conflict with the Preservation Criterion arises only when such appraisals relative to a belief state K are to be included in K itself. The way out suggested by Levi, is to ban such appraisals from the belief states with respect to which they are made and to confine them instead to 'higher-order' belief states. This move would allow to retain both (P) and a version of the Ramsey Test according to which a conditional $A > B$ ($\sim(A > B)$) is *true* with respect to a state of belief K whenever the minimal change of K needed to consistently include A does (does not) rule out *not-B* as a serious possibility.[2]

Gärdenfors is reluctant to adopt Levi's way out. The reason, as stated in (Gärdenfors 1986), is that there is no obvious way in which Levi can account for the acceptance conditions (or rather, truth conditions with respect to a state of belief) of iterated conditionals. But (Levi 1988) goes some way in alleviating such worries. It seems to me, however, that the introduction of a hierarchy of belief states is a rather desperate move in response to the conflict between (P) and the presence of conditionals or appraisals of serious possibility in belief states according to (RT) or its serious-possibility-analogue. It must be worth our while to explore the prospects for dispensing with such hierarchies.

Both Gärdenfors and Levi show a considerable determination to keep the Preservation Criterion (P). Yet, at the face of it, their incompatibility results say nothing more than that certain principles about rational belief change, notably (P), are sensitive to the underlying language and closure conditions on sets of sentences – representing states of belief – associated with certain connectives of this language. One should not be surprised that principles for rational change which are almost platitudinous when concerned with representations of belief states in a truth-functional language only, break down when the field of investigation is extended. It is only when such results are presented with an air of paradox that the move towards meta-. . . meta-belief states looks sufficiently motivated. In contrast, I shall argue that such a move is undermotivated when dealing with appraisals of serious possibility relative to states of belief. One should have no regrets to give up (P). I shall leave it as an open question whether abandoning (P) should also be recommended as the proper way to resolve the conflict between (P) and the Ramsey Test (RT).

This does of course not constitute an argument against the introduction of hierarchies of states of belief. There may very well be excellent reasons to introduce the kind of hierarchies proposed by Levi. But the incompatibilities mentioned above provide no sufficient reasons to proceed in this way. Indeed, there is a danger inherent in any quick move towards such hierarchies. Consider the case of subjunctive conditionals again. Suppose we agree that a subjunctive conditional $A > B$ should be construed as expressing a judgement about what will be accepted in a belief state revised to include A, i.e., a judgement about a potential transformation of a belief state K. Thus $A > B$ is true with respect to a belief state K whenever K has the right 'dispositional' property: that it would include B, if it were to be minimally changed to include A. Why should we not give the same status to the material conditional $A \supset B$? Should we not say that $A \supset B$ is true with respect to K whenever simply adding A to K commits to the acceptance of B? This is an appraisal with respect to a potential transformation of K just as much as the corresponding statement involving a subjunctive conditional. To be sure, we can agree that the material differs from the subjunctive conditional in many respects. But it does not suffice to just point out these differences unless we are also presented with an argument why at least some of those differences should be the grounds for assigning the subjunctive conditional a status so radically different from that of the material conditional in a theory of belief change.

In order to conduct the following discussion in a reasonably concise manner, we need to introduce a few conventions. I shall use the terms 'belief set', 'corpus', and 'theory' interchangeably. A belief set is a set of sentences in a countable sentential language containing a functionally complete set of truth-functional connectives. Belief sets are, until further notice, closed under truth-functional deducibility. Thus, the set of sentences truth-functionally deducible from a given set Γ, Cn (Γ) is defined, as follows:

(Cn)(a) $A \in \Gamma$ and $B \in \Gamma \Rightarrow A \& B \in Cn(\Gamma)$;

(b) $A \in \Gamma$ and $\vdash A \rightarrow B \Rightarrow B \in Cn(\Gamma)$;

(c) $Cn(\Gamma)$ is the smallest set satisfying conditions (a)–(b).

The turnstile signals theoremhood in the truth-functional logic TV. By defining *Cn* in terms of TV, I do not intend to reveal any preference for classical logic. It will just facilitate the discussion not to deviate unnecessarily from the assumptions shared by Gärdenfors, Levi, and many others who have written on the topic of theory change. I shall show later that nothing much hinges for our concerns in this paper on the choice of logic for the definition of the consequence operation (unless, of course, one chooses an excessively weak logic).

It should be noted that belief sets in the sense defined in the last paragraph are not intended as formal representations of states of belief in any intuitive sense. The belief states of ordinary agents are not closed under logical consequence. Levi's term 'corpus' is free from such unintended connotations. The corpus of an agent comprises all propositions he is committed to *qua* his avowed beliefs and qua logical consequence – regardless of whether the agent does or even can live up to his commitments. If the reader doubts whether there is the kind of systematic connection between ordinary agents' beliefs and corpora indicated in the last sentence, then I recommend to read 'theory' for 'corpus' or 'belief set'.

In Section 2, I shall present the 'paradox of serious possibility'. The paradox consists of a few *prima facie* plausible assumptions from which a contradiction is derived by means of commonly accepted principles of inference. The plausible assumptions are principles concerning rational belief (or theory) change, and for those readers who have not yet given the subject of theory change much thought, I shall try to make all principles involved as plausible as I can.

In Section 3, I shall consider five different responses to the paradox. The response I favour requires giving up the Levi Identity which says that revising K to include A is the same as first retracting $\sim A$ from K and then adding A. But Levi has argued that if revisions are to be legitimate, then they have to be replaceable by some sequence of contractions and revisions (not necessarily in that order and not necessarily non-empty). In Section 4, I shall make a proposal as to how this requirement can be met without endorsing the Levi Identity.

2. THE PARADOX OF SERIOUS POSSIBILITY

The corpus of an agent is his standard of *serious possibility*. If agent X at time t accepts[3] a sentence A, then $\sim A$ is not a serious possibility for X at t. Also, if X does not accept A at t, then $\sim A$ is not ruled out by his corpus; so $\sim A$ is a serious possibility for X at t. Suppose Xaver knows that Yvonne believes that A. Then Xaver knows that $\sim A$ is not a serious possibility relative to Yvonne's corpus of knowledge. If Yvonne believes what Xaver tells her, then Yvonne will believe that $\sim A$ is not a serious possibility according to her own standard of serious possibility. I use M_{K_Y} as a serious possibility operator relativised to the corpus K_Y. Let K_X stand for Xaver's corpus of knowledge. Similarly, let K_Y stand for Yvonne's corpus of knowledge. $M_{K_Y}A$ may be read as 'Given what Yvonne accepts as true, A is a serious possibility for her'. The above train of thought may then be represented as follows.

$$A \in K_Y \Rightarrow \sim M_{K_Y} \sim A \in K_X \Rightarrow \sim M_{K_Y} \sim A \in K_Y.$$

In expressions of the latter kind we may conveniently drop the subscripts, fixing contextually the relativity of the M-operator to a corpus and the relativity of corpora to agents.

It is uncontentious that, when talking *about* a particular corpus, we can say what, according to that corpus is or is not a serious possibility. That is to say, from any given corpus we can generate a meta-corpus Poss(K) according to the following rule.

(Poss)(a) If $A \in K$, then $\sim M \sim A \in \text{Poss}(K)$.
(b) if $A \notin K$, then $M \sim A \in \text{Poss}(K)$.
(c) Poss(K) is the smallest set satisfying conditions (a)–(b).

The considerations of the last paragraph suggest that there is a natural route from making judgements of serious possibility about corpora to having such judgements as members of corpora. At least there is no immediate reason why it should be ruled out in advance that judgements of serious possibility can make their appearance in belief sets. That is, there is no immediate reason why belief sets should not include sentences formulated in a language L^m which includes the operator M in its set of primitive symbols.

(O) Belief sets may include any well-formed sentence of the language L^m.

Yvonne took a detour in acknowledging that $\sim A$ is not a serious possibility according to her own standards. Such a detour would be unnecessary, if Yvonne were an ideal agent, recognizing all of her commitments and all of her suspensions of judgements. For then she would have access to a complete standard of serious possibility, i.e., one that yields a yes or no answer for each sentence of the language as to its serious possibility. Thus Yvonne would incorporate judgements of serious possibility in to her corpus K according to the definition of the Poss operation plus the additional requirement that Poss(K) is to be integrated into K. A set Poss(K) of judgements of serious possibility is integrated into its standard K if and only if

(Poss) $\text{Poss}(K) \subseteq K$.

(Note that Poss – unlike Cn – is not a closure operation in the Kuratowski–Tarski sense. Although Poss is idempotent, i.e., it satisfies the condition

(Idem) $\text{Poss}(\text{Poss}(\Gamma)) = \text{Poss}(\Gamma)$,

both Monotonicity

(Mono) $\Delta \subseteq \Gamma \Rightarrow \text{Poss}(\Delta) \subseteq \text{Poss}(\Gamma)$,

and Inclusion

(Incl) $\Gamma \subseteq \text{Poss}(\Gamma)$

fail. Having said this, I shall go on saying that a set is 'closed under Poss'.)

Closing corpora under Poss has as an immediate consequence that corpora are *complete* with respect to judgements of serious possibility (M-complete): either $MA \in K$ or $\sim MA \in K$, for each atom or truth-functional compound A and each corpus K. But the requirement that corpora are completely opinionated as to serious possibility is different from the far less attractive requirement that corpora are opinionated with respect to *any* odd sentence of the language. We should want to deny this latter requirement. Thus

(IC) Corpora may be incomplete with respect to some sentences of the underlying language; i.e., not for all sentences A: either $A \in K$ or $\sim A \in K$.

Completeness in the sense of (IC) can only be achieved by an agent omniscient with respect to her past, present, and future environment. Completeness with respect to serious possibility does not require input from the environment past, present, or future. Similarly, closure under Poss is only misleadingly described as requiring that corpora are veridical with respect to serious possibility. Commitment to judgements of serious possibility according to a corpus K is a function of K alone. Real agents may fail to recognize what they should consider a serious possibility according to their professed beliefs for essentially the same reasons as those explaining why they sometimes fail to acknowledge the logical consequences of what they believe: limitations of memory and computational capacity. In contrast, agents fail to have only true beliefs for completely different reasons. Thus, once we have decided that corpora in L^m are theoretically interesting items to consider, closure under Poss will have the same methodological status as closure under Cn.

A belief change model M is represented by a quadruple $\langle \mathbf{K}, Cn, *, -\rangle$. $\mathbf{K}$ is a set of sets of sentences in the language L^m.[4] $K, K_1, K_2, \ldots$ will serve as variables ranging over $\mathbf{K}$. Members of $\mathbf{K}$ are subsets of L^m closed under Cn.

The contraction operation $-$ and the revision operation $*$ map members of $\mathbf{K}$ to members of $\mathbf{K}$. In the literature a number of conditions on $-$ and $*$ have been proposed.[5] Such conditions are meant to capture the following intuitive conceptions of two kinds of theory change.

> The *contraction* of a corpus K by a sentence A should be a corpus $K-A$ obtained from K by removing as little from K as required for not having A among its members.
>
> The *revision* of a corpus K by a sentence A should be a corpus $K*A$ obtained from K by changing as little in A as required for adding A consistently to K.

We shall assume only four conditions on these change operations.

First, we shall assume that all members of $\mathbf{K}$ are closed under Poss. That is, we shall require that

(Poss) $\text{Poss}(K) \subseteq K$, for all $K \in \mathbf{K}$.

A third kind of theory change can now be defined by means of the consequence operation Cn and the condition (Poss).

The expansion of a corpus K by a sentence A, $K+A$, is the deductive closure of $K \cup \{A\}$ satisfying (Poss). That is, $K+A$ is the smallest set such that

(a) $K \cup \{A\} \subseteq K+A$;
(b) if $B \in K+A$ and $C \in K+A$, then $B \,\&\, C \in K+A$;
(c) if $B \in K+A$ and $\vdash B \rightarrow C$, then $C \in K+A$;

(d) if $B \in K+A$, then $\sim M \sim B \in K+A$;
(e) if $B \notin K+A$, then $M \sim B \in K+A$.

Expansions, thus defined, will almost always – i.e., in all interesting cases – result in inconsistent sets. For, suppose that $A \notin K$, then, since K is a corpus and, hence, closed under Poss, $M \sim A \in K$, and so, by condition (a), $M \sim A \in K+A$. But $A \in K+A$; hence, $\sim M \sim A \in K+A$. This result makes expansions of corpora in L^m quite uninteresting. Yet, the definition of an expansion $K+A$ is an obvious extrapolation to corpora in L^m of the Alchourron-Gärdenfors-Makinson definition (which is just the deductive closure of $K \cup \{A\}$). In a sense, this is as it should be: expansions are meant to be simple-minded – they should just add data to a given corpus, not retracting anything. However, an assessment of the paradox of serious possibility bifurcates at this point. One may refuse to adopt our rather simple-minded definition of an expansion and define expansions as special cases of revisions instead. Hans Rott (Rott 1989) proposes just this and I have no decisive objection to his proposal. In this paper, another option will be pursued: keeping the simple definition of expansion, while showing how the trivialisation of the revision operation can be avoided.

Second, we shall assume that revision is a consistency preserving operation as far as Cn allows it to be one. If K is consistent and $\sim A$ is not in every member of **K** in virtue of its being a theorem of the logic underlying *Cn*, then the set K revised to include A should be consistent. Let Λ denote an arbitrary inconsistent set. Since we have defined Cn in terms of truth-functional consequence, it follows by *ex falso quodlibet* that there is only one such set: the set of all sentences well-formed in L^m

(C) $K \neq \Lambda$ and $\nvdash \sim A \Rightarrow K^*A \neq \Lambda$.

Third, we shall assume that if a sentence A is not in a corpus K, then contracting K by A is vacuous: removing something from a set which is not there, is an empty gesture; it does not change anything.

(V) $A \notin K \Rightarrow K - A = K$.

Fourth, we shall assume that every revision must be decomposable into a contraction followed by an expansion. If we want to revise K by A, then we first clear the way by contracting $\sim A$ and then we just expand $K - \sim A$ by A. The thesis that there is no justifiable primitive revision operation is due to Isaac Levi. Thus the following condition on belief change models will be called 'the Levi Identity'.

(LI) $K^*A = K - {\sim}A + A$.

We now have gathered all data for presenting the main result of this section.

THEOREM 1. No belief change model can satisfy all of the conditions (IC), (Poss), (V), (C), and (LI).

Proof. We envisage a corpus K which, in accordance with (IC), is not opinionated with respect to some sentence A.

(1)	$A \notin K$,	Assumption
(2)	${\sim}A \notin K$,	Assumption
(3)	$M{\sim}A \in K$,	(1), (Poss)
(4)	$K = K - {\sim}A$,	(2), (V)
(5)	$K + A = K - {\sim}A + A$,	(4)
(6)	$K + A = K^*A$,	(5), (LI)
(7)	$A \in K + A$,	Definition of +
(8)	${\sim}M{\sim}A \in K + A$,	(7), (Poss)
(9)	$K \subseteq K + A$,	Definition of +
(10)	$M{\sim}A \in K + A$,	(3), (9)
(11)	$K^*A = \Lambda$,	(6), (8), (10)
(12)	$\nvdash {\sim}A$,	(2)
(13)	$K \neq \Lambda$,	(2)
(14)	$K^*A \neq \Lambda$,	(12), (13), (C)
(15)	*Contradiction!*,	(11), (14). ∎

A closely related result can be obtained by replacing (V) and (LI) in Theorem 1 by the Preservation Criterion (P) and the 'success' condition

(S*) $A \in K^*A$.

Proof. Assume (1) $A \notin K$ and (2) ${\sim}A \notin K$. Then, from (1) by (Poss), (3) $M{\sim}A \in K$, and, from (2) by (P), (4) $K \subseteq K^*A$; hence, (5) $M{\sim}A \in K^*A$. (6) $A \in K^*A$ is an instance of (S^*), and so, by (Poss), ${\sim}M{\sim}A \in K^*A$. But then, putting together (5) and (7), (8) $K^*A = \Lambda$. From (2) we can infer both (9) $\nvdash {\sim}A$ and (10) $K \neq \Lambda$. Thus, using (C), it follows that K^*A is consistent, which contradicts (8).[6]

I shall not argue about (S*) but take it as a minimal condition on any revision operation. (P) is entailed by (V) and (LI). Thus (P) cannot be called into question without questioning either (V) or (LI). Since I shall argue that (IC) and (C) should not be abandoned and since it is the purpose of this paper to explore the consequences of keeping (Poss), we shall have to reject (P). This is the reason why I have preferred to formulate the incompatibility result directly in terms of the principles which underpin (P), viz. (V) and (LI).

I have tried to make as strong a case as I could for each of the principles involved in the foregoing proof. Depending on how well I have succeeded in this effort, you will find the result reported in the theorem more or less paradoxical. In the next section I shall discuss various ways out of the paradox.

3. RESPONSES TO THE PARADOX

To avoid the contradiction in line (15) of the proof, we have to reject the derivation of either (11) or (14).

1. There is little we can do about (14). Given the definition of the consequence operation, (14) follows immediately from our only two assumptions expressing that K is not opinionated with respect to A. We may of course want to define the consequence operation by means of a weaker logic, say one of the family of relevant logics. In a relevant logic we would neither have the principle *ex falso quodlibet*,

(EFQ) $A \mathbin{\&} {\sim}A \rightarrow B$,

nor would closure under consequence include automatically all theorems of the logic in question in a corpus. Accordingly, neither could we infer from $A \notin K$ that K is consistent, nor could we infer from ${\sim}A \notin K$ that ${\sim}A$ is not a theorem of the underlying logic. But the gist of the principle (C) was this: if K is consistent so should K^*A. Given the intended interpretation of '*', the consistency principle sounds eminently plausible though its particular formulation in (C) may have too much of an eye on the vagaries of truth-functional logic. If '$\neq \Lambda$' is interpreted as a predicate of sets, expressing the property of consistency, then we would still want

(CC) $K \neq \Lambda \Rightarrow K^*A \neq \Lambda$.

We surely also want **K** to include consistent corpora. But if we assume

K to be one of these consistent corpora in **K** we have (14) back by detaching the consequent from (CC).

It should, incidentally, be noted that in the course of the proof a result with a somewhat paradoxical flavour has already been derived in line (11): not only can we not consistently add A to a corpus consistent with A (i.e., not including $\sim A$), but neither can we revise such a corpus to consistently include A. It is thus line (11), not line (15), which we should want to reject in order to avoid a contradiction.

2. Worries about the principle (IC) should not delay us for a minute. Giving up (IC) amounts to restricting the applicability of belief change models to maximal corpora: maximally consistent corpora or otherwise, given the classical consequence operation, maximal – i.e., trivial – corpora. If this were the only way out of the paradox, it could justly be considered a trivialisation result.

3a. The only serious candidates to blame are (Poss), (V), or (LI). Levi rejects (Poss).[7] In fact, Levi rejects (O), the principle that allows corpora in the universe set of a belief change model to include sentences expressing judgements of serious possibility side by side with non-modal sentences. In support of his rejection of (O), Levi points out a number of characteristics that set judgements of serious possibility apart from other kinds of judgements, the chief difference being that judgements of serious possibility belong to a class of judgements that are *about* corpora. But why should judgements about corpora not figure within corpora? The case of Yvonne and Xaver is a straightforward example of how judgements of serious possibility may enter corpora in essentially the same way as any other judgements. A case has yet to be made for why judgements of serious possibility with respect to a corpus K have to be confined to a *meta*-corpus Poss(K). Curiously enough, it is essentially the argument presented above that Levi takes to provide the required case against (O).[8] It will be the task of the remainder of this section to show that, while (O) continues to stand as a firm intuition, one of Levi's most cherished principles, namely (LI), is the most plausible principle to blame for the paradoxical result.

3b. Another option would be to accept (O) and reject (Poss) directly. Thus, it may be conceded – *pace* Levi – that there is no principal obstacle to having judgements of serious possibility in belief sets. But the completeness requirement with respect to such judgements – due to the systematic way in which they are to be included according to (Poss) – should be rejected. I have already argued above that the completeness requirement with respect to serious possibility should not be confused with the negation of (IC). For belief sets in a truth-functional language L, closure under logical consequence reflects our theoretical interest in corpora which are complete with respect to commitment. The least we can do to reflect that interest with respect to corpora in L^m, is to close them under Poss. (Perhaps we should do more: close them under some modal logic as well.) But we do not even need the full force of (Poss), requiring corpora to be 'veridical' with respect to serious possibility. We only need the extra, very weak, assumption that at least some corpora include judgements of serious possibility which are correct according to the rules provided by (Poss). That is, we weaken the theorem as follows. Call a corpus K *weakly M-*veridical if and only if there is at least one sentence A in the underlying language such that $M{\sim}A \in K$ if $A \notin K$ and ${\sim}M{\sim}A \in K$ if $A \in K$.

THEOREM 2. Consider a belief change model M, satisfying (O), (IC), (V), and (LI). Let K be a corpus in M and assume

(a) K is not opinionated with respect to A,
(b) K and $K+A$ are weakly M-veridical with respect to A.

Then (i) K can not be consistently revised to include A, and (ii) if M also satisfies (C), then M is contradictory.

Proof. The proof of (ii) is just like the proof of Theorem 1, except that (3) and (8) are now assumptions. The proof of (i) is obtained from the proof of (ii) by deleting lines (12)–(15). ■

What the theorem says may be put thus: any corpus that gets its judgements of serious possibility right with respect to at least one sentence on which it suspends judgement, will be 'punished' – it either has to suspend judgement forever or revise into inconsistency; if we also assume (C), contradiction ensues again.

4. The condition

(V) $\quad A \notin K \Rightarrow K - A = K,$

looks too much like a platitude to be seriously considered for rejection. Yet, when applied to corpora in L^m, (V) has the consequence that sometimes contraction does not do what we should expect it to do. We have not said much about the precise properties a change operation should have in order to qualify as a contraction operation. In particular, we have not provided an explicit recipe for how to construct from a given corpus K and a given sentence A the contraction of K by A, $K - A$. But our informal characterization in Section II, suggested that contractions should incur a *minimal loss of information*. If K is to be contracted by A, then we should remove from K only as much as is strictly necessary for retracting A. But there is another intuitive constraint which derives from the purpose for which we, at least in some cases, want to contract a corpus: contractions should be mind-opening. Sometimes, when we contract by $\sim A$, we do this with a view to take A into the corpus – we want to clear the way for consistently adding A. Indeed, (LI) (in conjunction with (C)) is perhaps best understood as requiring that all contractions be mind-opening. For suppose that K is consistent and that $\sim A$ is not a logical truth. Then, if $K - \sim A$ is not a mind-opening contraction, then $K - \sim A + A$ cannot be the same as K^*A, as asserted by (LI), given that * has to satisfy condition (C). Thus, if (LI) is to hold, then the contraction operation involved in (LI) has to be a mind-opening one.

There is no reason to suppose that for belief change models in a purely truth-functional language, the requirements that contraction should incur minimal loss and be mind-opening could come into conflict with each other. But the result recorded in Theorem 1 may be interpreted as showing that such a conflict can emerge when we extend our considerations to belief change models in L^m. Step (4) in the proof of Theorem 1,

(4) $\quad K = K - \sim A,$

may be identified as the step that foreshadows disaster. For, if (4) is to be accepted, then $K - \sim A$ does not allow to consistently add A: $M \sim A \in K - \sim A$ and so both $M \sim A$ and $\sim M \sim A$ are in $K - \sim A + A$. But if contracting K by $\sim A$ should allow to consistently add A, then $M \sim A$ has to be removed from K. But then $K - \sim A$ has to be a proper subset of K, contrary to (V). To conclude, if (LI) is to hold, then the contraction operation figuring in (LI) has to be a mind-opening one. But for corpora in the language L^m, (V) fails for mind-opening contraction. (V) may still hold for minimal-loss contraction. But minimal-loss contraction is not the kind of operation for which (LI) should be expected to be true. Thus the proof of Theorem 1 is flawed by an equivocation of two kinds of contraction operations. These operations happen to coincide in all their properties – or so we have reason to suppose – as long as we confine our attention to a rather simple language. But extending our theory to a language like L^m presses upon us the need to make an important distinction.

I shall only indicate here how the diagnosis of the last paragraph can be turned into a refinement of the belief change models discussed so far. A *mind-opening belief change model* is a sextuple $\langle \mathbf{U}, \mathbf{K}, \mathbf{N}, Cn, \dot{-}, \hat{-} \rangle$. $\dot{-}$ is meant to stand for the mind-opening contraction operation; $\hat{-}$ is meant to stand for the minimal-loss contraction operation. **U** is a set of subsets of L^m closed under Cn. **K** is the set of all proper corpora, i.e., all those sets in **U** satisfying (Poss). All members of **U** not satisfying (Poss) (improper corpora) are collected in **N**. Thus

(a) $\mathbf{U} = \mathbf{K} \cup \mathbf{N}$;
(b) $Cn\colon \mathbf{U} \to \mathbf{U}$;

The expansion operation + is defined as before: $K + A$ is the deductive closure of $K \cup \{A\}$ satisfying (Poss), for all $K \in \mathbf{U}$. Thus + maps any member of $\mathbf{U} \times P(L^m)$ into a proper corpus, i.e., a member of **K**. Minimal-loss contraction maps any pair consisting of a member of **U** and a sentence into a member of **K**, i.e., a proper corpus;

(c) $\hat{-} : \mathbf{U} \times L^m \to \mathbf{N}$.

Mind-opening contraction maps any pair in $\mathbf{U} \times L^m$ into a member of **N**;

(d) $\dot{-} : \mathbf{U} \times L^m \to \mathbf{K}$.

We require (LI) and (C) for $\dot{-}$ and (V) for $\hat{-}$. There are other conditions on $\dot{-}$ and $\hat{-}$ which are plausible to add, but they need not interest us now.

The reason for introducing the set **N** into the models is that we do not want sets that have been mind-openingly contracted to be closed under (Poss) again – otherwise (V) would hold for $\dot{-}$ after all and (LI) could not be a condition on $\dot{-}$, on pain of contradiction. Thus $\dot{-}$ has to be an operation that does not yield proper corpora. Nevertheless, a revised set K^*A will always be a proper corpus in virtue of (LI) and the definition of +.

For Levi, deductively closed sets are epistemologically interesting because they provide a standard for serious possibility; given a deductively closed set K, it is determined for every sentence of the underlying language whether it is, or is not, a serious possibility for an agent committed to K. Earlier we expressed this fact by saying that every deductively closed set K gives rise to an M-complete set of judgements of serious possibility relative to K. We have argued so far that – *pace* Levi – there is no good reason for keeping K and Poss(K) apart. We have argued that K and Poss(K) should be integrated to a single unit of inquiry, a corpus closed under both logical consequence and serious possibility. Accordingly, if rational inquiry is to proceed by rational change of corpora, then rational changes should satisfy the requirement that they map one M-complete corpus into another. In mind-opening belief change models the mind-opening contraction operation fails to deliver M-complete corpora. Hence, in these models, $\dot{-}$ is not a rational change operation. Mind-openingly contracted sets have to be M-incomplete, which is just another way of saying that either they should play no part in a normative epistemology or they illustrate that, at least sometimes, a standard of serious possibility has to be kept apart from the serious possibility judgements it gives rise to. Both of these alternative conclusions suggest that mind-opening belief change models cannot serve their intended purpose.

5. The discussion of mind-opening belief change models proceeded on the assumption that we wanted to satisfy Levi's requirement that revisions are to be analysed as sequences of contractions and expansions. That is, we wanted to retain the Levi Identity

(LI) $K^*A = K - {\sim}A + A,$

for some kind of contraction operation. It turned out that the contraction operation involved had to be a mind-opening one, not satisfying (V). But such a contraction operation cannot be modelled without introducing improper corpora. If, for the one or the other reason, we do not like to have improper corpora in our belief change models, then we will have to do without mind-opening contraction. Yet, if the contraction in (LI) is not mind-opening, then we have no reason at all to suppose that the Levi Identity should be true. Furthermore, giving up (LI) allows us to hold on to (V) which, given that – is not required to be mind-opening, is an attractive principle to have. Unless we have good independent reasons for believing that (LI) is absolutely indispensable, we should hold on to the view, expressed in (O), that judgements of serious possibility do not require a separation of beliefs into corpora of different levels, while being attentive to the fact that some of the principles concerning belief change are sensitive to the language in which the units of change are formulated.

4. REVISION DECOMPOSED

Let me briefly summarize the result of the discussion in the foregoing section. In response to the paradox of serious possibility the following *consistent* picture has emerged. Belief sets need not be opinionated on every sentence of the underlying language. Yet they are completely opinionated with respect to what is, or what is not, a serious possibility. Every belief set contains such a *complete* set of serious possibility judgements. We can allow for vacuous contraction: if A is not in a certain belief set K, then nothing is removed from K when K is "contracted" by A. Furthermore, revision is a consistency preserving operation – after all, this is the purpose of revisions, as opposed to expansions. We have seen that in such a belief change model, the contraction operation cannot always be mind-opening. This is the reason why we feel entitled to give up (LI).

To those who share Levi's epistemological tenets, abandoning (LI) may seem enough to cast a dark shadow on the picture drawn in the last paragraph. In barest outline, the reason is this.

Levi has forcefully argued for a thesis which he calls *epistemological infallibilism*: no matter how a piece of information has entered the corpus of beliefs of an agent (i.e., no matter what its pedigree), once an agent assumes a proposition as part of his or her corpus, the agent has to treat it as infallibly true.[9] This view has an at first startling but simple consequence for the concept of revision:

> From X's initial point of view when he adopts a theory T_1, replacement of T_1 with T_2 inconsistent with T_1 is tantamount to the deliberate substitution of a certainly false hypothesis for one which is certainly true. [. . .] If [. . .] X intends to avoid error in revising his corpus, it is utterly counterproductive from his point of view to undertake the replacement. [. . .]
>
> The legitimacy of replacement could be denied; but this would be to condemn a historically important mode of development of scientific knowledge as illegitimate. [. . .]
>
> There is, however a[n] [. . .] alternative. We might deny the legitimacy of efforts to justify shifting directly from T_1 containing h to T_2 containing $\sim h$, but might allow the legitimacy of such replacement provided that it can be decomposed into a sequence of contractions and expansions each of which is justified.[10]

Thus, according to Levi, a revision (replacement) cannot be justified unless it can be decomposed into justifiable steps of contraction and expansion. We do not need to concern ourselves here with the question as to the conditions under which contractions, respectively expansions, can be justified. (LI) certainly is *one* way of giving formal expression to the Decomposition Principle, i.e., (DP) Every legitimate revision is decomposable into a sequence of contractions and expansions.

Yet rejection of (LI) does not entail rejection of (DP). I shall presently argue that the picture drawn in the first paragraph of this section can be consistently enriched with a surprisingly simple condition expressing that (DP) is satisfied. But, for obvious reasons, the new condition will not be equivalent to (LI).

First however, it will be worth our while to point out that the paradoxical result recorded in Theorem 1 is not particular to the notion of serious possibility. The reasoning that led to apparent paradox can be reinstantiated for any modality satisfying certain formal properties with regard to belief sets. A *reflective modality* (in a belief change model M) is any unary sentence forming operator (or compound of such operators) ϕ such that

(a) ϕ is not the identity function,
(b.1) $A \in K \Rightarrow \phi A \in K$ $(\forall K$ in M$)$,
(b.2) $A \notin K \Rightarrow {\sim}\phi A \in K$ $(\forall K$ in M$)$.[11]

It is easy to see, by inspection of the proof of Theorem 1, that any reflective modality will instantiate the paradox. For the paradox of serious possibility let $\phi = {\sim}M{\sim}$. ϕ may also be reasonably construed as a truth-predicate T (truth according to a corpus), as an assertion operator $\vdash$ (asserted according to a corpus), or, perhaps, as some subjective probability operator, etc.

What do we require of a corpus K^*A in a model including a reflective modality ϕ to count as the revision of a corpus K to consistently include A? The answer is already contained in the question.

(i) K^*A should be closed under the rules generating ϕ-sentences into corpora (i.e., the rules (b.1) and (b.2));
(ii) K^*A should include A;
(iii) K^*A should be consistent, provided ${\sim}A$ is not a logical truth and K is consistent (in accordance with (C)).

(i)–(iii) entail the following requirements:

(1) $A \in K^*A$, (ii)
(2) $\phi A \in K^*A$, (1), (i)
(3) ${\sim}\phi A \notin K^*A$, (2), (iii)
(4) ${\sim}A \notin K^*A$, (1), (iii)

(5) $\sim\phi\sim A \in K^*A$, (4), (i)
(6) $\phi\sim A \notin K^*A$, (5), (iii).

The task is now to find a set H which results from K by use of the contraction or expansion operation alone and which can replace K^*A in (1)–(6) *salva veritate*. H then may reasonably be claimed to be a representation of the revision of K by A.

Let H be $K - \sim\phi A$. To verify (1)–(6) with $K - \sim\phi A$ in place of K^*A we need, apart from closure of $K - \sim\phi A$ under (b.1) and (b.2), only one additional principle, namely

(S-) $K \neq \Lambda$ and $\nvdash A \Rightarrow A \notin K - A$.

In words: contractions are successful, provided they are made on a consistent set and provided that what is to be contracted is not a logical truth. In showing that $K - \sim\phi A$ can replace K^*A in (1)–(6) I shall assume that the antecedent condition of (S-) is fulfilled. In virtue of (C), the equation

(*) $K^*A = K - \sim\phi A$,

holds trivially whenever $\sim A$ is a logical truth or K is inconsistent.

(3) $\sim\phi A \notin K - \sim\phi A$, (S), initial assumptions
(1) $A \in K - \sim\phi A$, (3), (b.2) contraposed
(2) $\phi A \in K - \sim\phi A$, (1), (b.1)
(4) $\sim A \notin K - \sim\phi A$, (1), $K - \sim\phi A \neq \Lambda$
(5) $\sim\phi\sim A \in K - \sim\phi A$, (4), (b.2)
(6) $\phi\sim A \notin K - \sim\phi A$ (5), $K - \sim\phi A \neq \Lambda$.

Note that no information is needed as to whether A or $\sim A$ is or is not in K. Thus in particular, $K - \sim\phi A$ may represent K^*A when K is not opinionated with respect to A, which was one of the conditions needed in Theorem 1.

Let K^*A be defined as $K - \sim\phi A$, then the following principles about revision can easily be verified on the assumptions made in this section; some of these principles are in fact identical with assumptions already made.

(*1) K is a corpus $\Rightarrow K^*A$ is a corpus.
(*2) $A \in K^*A$
(*3) $K^*A \subseteq K + A$
[*4] $\sim\phi A \notin K \Rightarrow K + A \subseteq K^*A$

(*5) $K \neq \Lambda$ and $\nvdash \sim A \Rightarrow K^*A \neq \Lambda$

(*6) $\vdash A \leftrightarrow B \Rightarrow K^*A = K^*B$.

Readers familiar with the work of Alchourron, Gärdenfors, and Makinson on theory change will, with the exception of [*4], recognize in (*1)–(*6) a complete list of the basic Gärdenfors postulates for revision. [*4] replaces Gärdenfors's

(*4) $\sim A \notin K \Rightarrow K + A \subseteq K^*A$,

which fails on the above definition: suppose $A \notin K$; then $\sim \phi A$ will be in K and also in $K + A$. Yet in all but trivial cases $\sim \phi A$ will not be in K^*A. Thus, for belief change models including a reflective modality ϕ, I propose $K - \sim \phi A$ as a representation of K^*A, thereby rounding off the picture drawn in the first paragraph of this section to satisfy the Decomposition Principle (DP).

NOTES

* I am grateful to Peter Lavers and Isaac Levi for the discussions that led to writing this paper. I also wish to thank Peter Gärdenfors and Hans Rott for their comments on an earlier version of this paper. This paper was written while holding a research scholarship in the Automated Reasoning Project at the Australian National University.

1 Levi (1974), p. 157.

2 See Levi (1988), Section V.

3 I use 'accepts' from now on as short for 'is committed to accept'.

4 By abuse of notation, I use L^m to denote both the structure specifying the language and the set of sentences well-formed according to the formation rules of the language.

5 For an informal survey of the theory of theory change see Makinson (1985). For the technical details see Alchourron et al., (1985).

6 The argument is more akin to Levi's in Levi (1988).

7 See Levi (1978, 1980, 1988).

8 The argument appears in Levi (1988). Levi's argument is the source of the theorem in the preceding section.

9 For a detailed statement of this view see Levi (1980).

10 Levi (1980).

11 Condition (a) is needed to rule out that (b.2) may require corpora to be negation-complete.

REFERENCES

Alchourron, C., P. Gärdenfors, and D. Makinson: 1985, 'On the Logic of Theory Change: Partial Meet Contraction and Revision Functions', *JSL* **50**, 510–30.

Gärdenfors, P.: 1986, 'Belief Revision and the Ramsey Test for Conditionals', *Philosophical Review* **95**, 81–93.
Levi, I.: 1984, 'Serious Possibility', in *Decisions and Revisions*, Cambridge University Press, Cambridge, pp. 147–61.
Levi, I.: 1980, *The Enterprise of Knowledge*, MIT Press, Cambridge, Massachusetts.
Levi, I.: 1988, 'Iteration of Conditionals and the Ramsey Test', *Synthese* **76**, 49–81.
Makinson, D.: 1985, 'How to Give It Up: A Survey of Some Formal Aspects of the Logic of Theory Change', *Synthese* **62**, 347–63.
Rott, H.: 1989, 'Conditionals and Theory Change: Revisions, Expansions, and Additions', *Synthese* **81**, this issue, 91–113.

Automated Reasoning Project
and Department of Philosophy
Research School of Social Sciences
Australian National University
GPO Box 4
ACT 2601
Australia