

A Tale of two Schemata: Tarskian (Finitary) Truth and Ramseyan Mental States

Arnold Koslow

“Our task, then, is to elucidate the terms true and false as applied to mental states, and as typical of the states with which we are concerned we may take for the moment beliefs.”¹

1.

I want to connect up into a coherent theory three primitive ideas which no one in their right mind would normally want to consider together. The first is this: In an (1922), unpublished (seriously unpublished) typescript on Truth and Simplicity, Frank Ramsey speculated on the idea that truth is an *incomplete symbol* and that the claim that “p’ is true” can be expressed in certain linguistic contexts by adding on the phrase “and p” in that context. As we shall see, that can’t be right, but, as we shall also see, that insight is part of a more general account of truth for finite languages introduced by Tarski. The second is also an insight of Ramsey about belief, which he called a “truism”. It can be expressed by saying that “Richard believes that p” is true if and only if “Richard believes that p”, and p (we use “p” as a schematic letter). We shall see, that can’t be right either, but it will become part of an account that we shall give that might be right. The third idea of Ramsey, was a central idea in an unpublished four chapter manuscript on logic; a late writing. The proposal was to organize an account of logic in terms of the truth and falsity of *belief states* rather than sentences, statements, or propositions. It was a very bold idea for that time and ours as well. That account never emerged, and it is these three ideas that I will try to combine into a coherent, simple theory.

Ramsey, in a remarkable passage of his unpublished *Facts and Propositions*, (1927), proposed a thesis about belief according to which (**RBT**) any belief that p, is true if and only if p. Here is that prescient passage:

“It is, perhaps also immediately obvious that if we have analyzed judgment we have solved the problem of truth; for taking the mental fac-

¹ “The Nature of Truth” in Rescher and Majer (1991, Chapter 1, p. 7).

tor in a judgment (which is often itself called a judgment), the truth or falsity of this depends only on what proposition it is that is judged, and what we have to explain is the meaning of saying that the judgment is a judgment that a has R to b , i.e. is true if aRb , false if not. ... In order to proceed further, we must now consider the mental factors in a belief.”²

This doesn't seem to be right when one thinks of it as a thesis about beliefs. However, if we adjust Ramsey's proposal so that it is a requirement on belief states, then it yields what I shall call *Ramsey's Belief Schema*, (**RBS**).³

(**RBS**). A belief state that p , is a true belief state, if and only if p .

Given our account of the truth and falsity of belief states, we will be able to prove this schema and it will be evident that it is a parallel that it is a close cousin of the Tarski T-schema. We turn first to a consideration of Tarski's definition of truth for finite sets of sentences

2.

Astute scholars of Ramsey and Tarski have noted that there is some connection, perhaps even an anticipation of Tarski's T-Schema by Ramsey's endorsement of what is sometimes called the *Redundancy Theory of Truth*. Truth be told, something akin to the T-schema goes back to Aristotle, resonates to Frege, and was endorsed by the early twentieth century Cambridge philosopher W. E. Johnson as a redundancy claim that he expressed in a form that is tied to assertions:

“the assertion of p is equivalent to the assertion that p is true ...”⁴

This is close to Ramsey's formulation but of course is not to be identified with the T-condition as formulated by Tarski.

There is however, another claim about truth and belief that Ramsey regarded as a “truism”, which in its own way appeals to another kind of redundancy. He describes the idea this way:

(**BTp**) A belief is true if it is a belief that p , and p

² “Facts and Propositions” in Mellor (1990, p. 39-40).

³ (**RBS**) will be expressed below by using schematic letters rather than propositional variables, and it is, as we shall see, analogous to the Tarski T-schema. This is a departure from Ramsey's way of expressing the condition, but not of any significance for the present discussion.

⁴ This history is succinctly and aptly documented in Sahlin (1990, Chapter 2).

We know from the published and unpublished papers of Ramsey that although he termed it a “truism”, it was nevertheless a central element of his thought about judgements, assertions or beliefs.⁵ It is also the sort of conjunctive construction that Ramsey made much of in an earlier unpublished talk which he gave to the Apostles in 1922.⁶

It is obvious that there are several important differences that should be noted. First, that the T-schema as it figured in Tarski’s work, and the variants of it that figure in the earlier accounts from Frege to W.E. Johnson are very different from each other, and second, that Ramsey’s version of the T-schema and his other leading dictum about true belief (**Btp**) are important but patently different. The difference is that Ramsey’s version of the T-schema leans heavily on epistemic notion of assertion or belief, while Tarski’s version involves sentences, and is well known for its eschewal of any reference to assertion, judgement, or belief. In light of these obvious differences, it would be plausible to conclude that there is no connection whatever between Ramsey’s version of the T-schema and his doxastic truism, and even less connection between Tarski’s version of the T-schema and that truism. Nevertheless, I want to argue that there is a way of showing how two basic ideas: Ramsey’s idea that the truth of *p* can be regarded in certain contexts as the conjunction of “and *p*” to that context, and his endorsement of the doxastic truism, can be brought into a very simple framework when they are both placed against the background of the Tarski T-schema in a *finitary* context. Consequently, when we look at that famous paragraph in “Facts and Propositions” where Ramsey says that

“It is, perhaps also immediately obvious that if we have analyzed judgment we have solved the problem of truth; for taking the mental factor in a judgment (which is often itself called a judgment), the truth or falsity of this depends only on what proposition it is that is judged, and what we have to explain is the meaning of saying that the judgment is a judgment that *a* has *R* to *b*, i.e. is true if *aRb*, false if not. ... In order to proceed further, we must now consider the mental factors in a belief.”⁷,

we shall see how so much of it falls easily into place –including an argument for the thesis that if there is a belief state that *p*, then it is a true belief state if and only if *p* –which we shall call *Ramsey’s Belief Thesis (RBT)*.⁸

⁵ This point is deftly made by U. Majer (1991).

⁶ Ramsey proposed as early as 1922, that “*p* is true” is conveyed in certain linguistic contexts by using “and *p*” as a suffix in those contexts. Consequently he referred to “truth” as an incomplete symbol, similar to Russell’s definite descriptions and classes in the *Principia*. Cf. A. Koslow (2005).

⁷ “Facts and Propositions” in Mellor (1990, p. 39–40).

⁸ (**RBT**) will be expressed below by using schematic letters rather than propositional variables, and it is, as we shall see, analogous to the Tarski T-schema. This schematic version is a departure from Ramsey’s way of expressing the condition, but not of any significance for the present discussion.

3.

FiniteTarski. It is best to begin the discussion with the problem of defining “true” or better, the predicate “is true” in the case where we consider the case of defining such a predicate for a specified finite number of sentences. Tarski showed how this could be done using an appropriate metalanguage, without need for a notion of satisfaction. Before we consider his construction, it is instructive to look at some very simple special cases

Let’s begin with the simplest case: to define a truth predicate for the sentence A , such that it is provable that $\text{Tr}(|A|) \leftrightarrow A$, where any sentence flanked by two vertical lines will count as a name of that sentence. The definition is straightforward:

$$\text{Tr}(x) : (x = |A|) \wedge A.$$

To provide a proof that $\text{Tr}(|A|) \leftrightarrow A$, we assume that in the metalanguage we have $|A|$, and the axiom $|A| = |A|$. It is trivial to prove the equivalence of A and $\text{Tr}(A)$. First note that $\text{Tr}(|A|)$ is just $(|A| = |A|) \wedge A$, which implies A . For the converse, note that from A together with the axiom that $|A| = |A|$, we have their conjunction $(|A| = |A|) \wedge A$, which, by definition, is just $\text{Tr}(|A|)$. It was assumed in this argument, and in the ones to come that enough classical sentential logic is at hand to run the argument.

So for any single sentence we can define a predicate that is tailor-made to insure that it is a provable case of the Tarski T-schema. It is also obvious that for any specific sentence B , other than A , the sentence “ $\text{Tr}(|B|)$ ” is provably incorrect, provided that for the sentence B , we have the sentence $|B| \neq |A|$ as an axiom of the metalanguage. For then $\text{Tr}(|B|)$ is the conjunction $(|B| = |A|) \wedge A$, and the first conjunct contradicts one of the axioms of the metalanguage.

This construction, of a truth predicate, sentence-by-sentence, is not to everyone’s liking. True we have a schema such that for every A , $\text{Tr}(|A|) \leftrightarrow A$ is provable in the metalanguage, but it’s always a different predicate that is indexed to each sentence. It would be more precise to say that for every sentence A , the following instance of a truth-schema can be proved for a truth predicate that is indexed to A . i.e.

$$\text{Tr}_A(|A|) \leftrightarrow A.$$

The special feature of this schema is that it is doubly schematic. The “ A ” is a schematic letter, but the predicate is schematic too: there’s a *different* predicate “ Tr_A ” for each replacement of the schematic A . Now that may be a cause for complaint. It is normally assumed that the T-schema $\text{Tr}_A(|A|) \leftrightarrow A$ always uses the same predicate, no matter what particular sentence replaces the sche-

matic letter. One normally assumes that the sentences “‘A’ is true” and “‘B’ is true” use the same predicate. It is a familiar constraint on the various instances of the T-schema that is so common that it usually goes without saying. The doubly schematic version of the T-schema is highly relativized, but it is a coherent way of considering the relation between a statement A and the statement that A is true. It’s just not the usual thing.

There’s more dissatisfaction with this notion of truth that is indexed to single sentences: There seems to be no systematic relation between any of these truth predicates with each other, nor do any of them behave well with respect to logical operations. For example the truth of a conjunction does not imply the truth of its conjuncts, if truth is given by any of these statement-indexed predicates.

Many of these defects can be remedied by using a construction which Tarski employed in order to show, as he said, that “Under certain special assumptions the construction of a general definition of truth is easy”. He considered a finite fragment of English, or some object language with finitely many sentences.⁹ We shall use a slight modification which singles out some finite set of sentences $\Gamma_n = \{A_1, \dots, A_n\}$ either of a fragment of English or of some object language. The idea is that instead of defining a truth predicate for sentence A, and another for sentence B, that we define one truth predicate for both of them. In all the finite cases, that can be done disjunctively, as Tarski noted.

It is worth considering a few simple examples before describing his construction that covers the case of any finite set of sentences:¹⁰

(1) Let $\Gamma_1 = \{A, B\}$. Then set $T_{\Gamma_1}(x) = [(x = |A|) \wedge A] \vee [(x = |B|) \wedge B]$. We assume that the metalanguage contains $|A| = |A|$, $|B| = |B|$, and $|A| \neq |B|$ as axioms. Then it follows that $T_{\Gamma_1}(A) \leftrightarrow A$, and $T_{\Gamma_1}(B) \leftrightarrow B$ are provable in the metalanguage.

It is obvious that the mini-example of (1) will not guarantee that the truth of $(A \wedge B)$ will be equivalent to the conjunction of the truth of A and the truth of B. However if the conjunction of A with B is included in the set containing A and B, and the metalanguage has the self identities of $|A|$, $|B|$, $|(A \wedge B)|$, and the non-identities $|A| \neq |B|$, $|A| \neq |(A \wedge B)|$, and $|B| \neq |(A \wedge B)|$ in the metalanguage, then the result is easily obtained:

(2) Let $\Gamma_2 = \{A, B, (A \wedge B)\}$. Then set $T_{\Gamma_2}(x) = [(x = |A|) \wedge A] \vee [(x = |B|) \wedge B] \vee [(x = |(A \wedge B)|) \wedge (A \wedge B)]$. It follows that (i) $T_{\Gamma_2}(|(A \wedge B)|) \leftrightarrow (A \wedge B)$, (ii) $T_{\Gamma_2}(|A|) \leftrightarrow A$, and (iii) $T_{\Gamma_2}(|B|) \leftrightarrow B$ are provable in the metalanguage. Consequently, $T_{\Gamma_2}(|(A \wedge B)|) \leftrightarrow T_{\Gamma_2}(|A|) \wedge T_{\Gamma_2}(|B|)$ is also provable.

⁹ Tarski (1969, p. 63–77).

¹⁰ Following Tarski, we shall also assume that none of the sentences has occurrences of the predicate “true”. The aim after all is to define such a predicate.

As a final example we can consider how well this truth predicate behaves with respect to negation, if the negation of a sentence is included in the set of sentences for which the truth predicate is to be defined, and the metalanguage contains the self-identities of $|A|$, $|\neg A|$, and the non-identity $|A| \neq |\neg A|$.

(3) Suppose that $\Gamma_3 = \{A, \neg A\}$ and that $T_{\Gamma_3}(x) = [(x = |A|) \wedge A] \vee [(x = |\neg A|) \wedge \neg A]$. Thus (i) $T_{\Gamma_3}(|A|) \leftrightarrow A$, and (ii) $T_{\Gamma_3}(|\neg A|) \leftrightarrow \neg A$ are provable, and consequently, $T_{\Gamma_3}(|\neg A|) \leftrightarrow \neg T_{\Gamma_3}(|A|)$ is also provable. It follows from this that even in this simple case, $T_{\Gamma_3}(|A|) \vee T_{\Gamma_3}(|\neg A|)$ is provable (one of the conditions of adequacy that Tarski set for a definition of truth).

We can now proceed to Tarski's provision of a truth predicate for an arbitrary finite subset of the sentences of some object language, $\Gamma = \{A_1, A_2, \dots, A_n\}$:

$$T_{\Gamma}(x) : [(x = |A_1|) \wedge A_1] \vee \dots \vee [(x = |A_n|) \wedge A_n],$$

When the metalanguage contains as axioms the n identities $|A_1| = |A_1|$, ..., $|A_n| = |A_n|$, and all the non-identities $|A_i| \neq |A_j|$ (for all $i \neq j$ between 1 and n), it is easily proved that every instance, $T_{\Gamma}(|A_i|) \leftrightarrow A_i$ of the T-schema is provable in the metalanguage.

It is also interesting to note that if B is a consequence of the set $\Gamma = \{A_1, A_2, \dots, A_n\}$ but not a member of it, then one can form the larger set $\Gamma^* = \{A_1, A_2, \dots, A_n, B\}$, and for the truth predicate for the larger set, we will have preservation of the truth of a conclusion, given the truth of the premises: that is, if $A_1 \wedge A_2 \wedge \dots \wedge A_n \rightarrow B$ is a theorem of the metalanguage, then $T_{\Gamma^*}(A_1) \wedge T_{\Gamma^*}(A_2) \wedge \dots \wedge T_{\Gamma^*}(A_n) \rightarrow T_{\Gamma^*}(B)$ is also a theorem of the metalanguage (provided the usual assumptions of self identify and non-identity of the sentences of Γ^* are axioms of the metalanguage).

It is readily apparent that the truth predicates may differ with a difference of the sets for which they are defined. There is one fact which mitigates this feature, and it is that the truth predicate in the finite case is cumulative: If we suppose that Γ and Γ^* are two finite sets of sentences of the object language, such that one is a subset of the other, $\Gamma \subseteq \Gamma^*$, and " T_{Γ} " and " T_{Γ^*} " are their respective truth predicates, then in shifting from one set of sentences to a second larger one, even though there is a corresponding shift in the truth predicate, nevertheless, everything that is true with respect to the first predicate will also be true with respect to the second predicate. In this sense, truth is *cummulative*. Roughly stated: no truth's are lost in expanding the set of sentences for which truth is being defined. The reason lies mainly with the disjunctive character of the definitions of the truth predicate in the finite case. For example if Γ^* has only one more sentence say " B ", beyond the sentences of Γ , then $T_{\Gamma}(x)$ is $[(x = |A_1|) \wedge A_1] \vee \dots \vee [(x = |A_n|) \wedge A_n]$, and $T_{\Gamma^*}(x)$ is just $T_{\Gamma}(x) \vee [(x = |B|) \wedge B]$. So of course, it's provable that $[T_{\Gamma}(x) \rightarrow T_{\Gamma^*}(x)]$.

When the sentences to be covered are infinitely many, and where the structure of some of them involves quantification, the Tarskian construction proceeds differently from the finite case, using a more detailed theory concerning subsentential parts and the satisfaction relation. Reference becomes important, and the theory required begins to look more substantial than what was needed for the finite case.

4.

True Belief States. With these various familiar results for the notion of truth in the finite cases in place, it is plausible that even in such a simple case, a viable notion of truth can be provided that has many of the features that are so familiar. The Tarski construction for the finite case provides an account of the predicate “ x is true” ($T_{\Gamma}(x)$) for finite sets of sentences Γ of an object language essentially by exploiting a device that conjoins “and A” to a context indexed to A (“ $x = |A|$ ”) forming their conjunction (“ $[x = |A|] \wedge A$ ”), and then forming the disjunction of each such conjunction for each of the sentences of the set Γ . This part of the construction resonates with Ramsey’s early “conjunctive” view (1922) that “true” is an incomplete symbol. Our task now is to take to heart Ramsey’s proposal when, in his late study *The Nature of Truth*, he said that the

“... the task, then is to elucidate the terms true and false as applied to mental states, and as typical of the states with which we are concerned we may take for the moment beliefs.”¹¹

What we now wish to explain is how this task of elucidation of the truth or falsity of *mental states* might be accomplished by transferring or recreating in a parallel fashion, the Tarskian proposal for truth for finite sets of sentences to the nonlinguistic terrain of finite sets of mental states. One of the consequences of this elucidation is the proof for belief states of a counterpart of the T-schema for sentences, which we think is something that Ramsey anticipated.¹²

¹¹ N. Rescher and U. Majer (1991, p. 7).

¹² The focus on the problem of elucidating an account of the truth and falsity of mental states like belief states is explicit in the four unpublished chapters on logic probably written during 1927-8 according to R.B. Braithwaite (1931, p. xiii -xiv). It indicates a radical departure of course from the usual assumption that the predicates of truth and falsity apply to the usual suspects: sentences, statements, or propositions. In fact this focus on states of belief rather than the usual targets is what makes possible the development of the present account. It is not an account which Ramsey proposed. I agree with the judgment on those chapters given by Braithwaite who noted that Ramsey was profoundly dissatisfied with them, and accordingly did not include them in the 1931 collection.

Let us suppose that there are some special belief states, the belief state that p, the belief state that q, the belief state that r, It is for these kinds of mental acts for which a definition of "true" is intended. They are belief states with a special kind of content indicated by the use of "p", "q", "r", Ramsey, in speaking of these mental states says that

"... whether or not it is philosophically correct to say that they have propositions as objects, beliefs undoubtedly have a characteristic which I make bold to call *propositional reference*. A belief is necessarily a belief that something or other is so-and-so, for instance that the earth is flat; and it is this aspect of it, its being "that the earth is flat" that I propose to call its propositional reference."¹³

It is unimportant how this is decided, as long as they stand for the kinds of things that can enter into logical relations. Furthermore, just as in the Tarski construction, we needed some standard way to refer to the sentences under consideration, so too we shall use "tbtp", "tbtpq", "tbtr", ... as standard names to refer to the belief states: the belief state that p, the belief state that q, the belief state that r,

When we shall consider sets of tbtps, I shall assume that there is some one individual¹⁴ whose belief states we are considering. The individual needn't be a person, it could be one of Ramsey's wonderful creations - a chicken who believes that a certain sort of caterpillar is poisonous.¹⁵

We assume that for these mental states the belief state that p is identical with the belief state that p, and similarly for the belief state that q, and so on. That is, tbtp = tbtp, tbtpq = tbtpq, tbtr = tbtr, etc. Furthermore, if we have a set of belief states {tbtp, tbtpq, ... , tbtr} we shall assume that they are all different: that is, tbtp ≠ tbtpq, tbtp ≠ tbtr, tbtpq ≠ tbtr, etc.

We shall not assume any special way in which the states of belief such as tbtp, are related to their propositional references. For example we shall see that

¹³ "The Nature of Truth" in Rescher and Majer (1991, p. 7). The notion of propositional reference was explicitly described by Ramsey as primitive. However those beliefs that have propositional reference are those beliefs which are that something or other is a so-and-so. The simplest way to understand his view is that he is restricting his discussion to "belief that ..."; only he thinks that all beliefs are beliefs that. It should be said that there is some deliberate vagueness as well in Ramsey's use of the notion of "state", but he does not include sentences, statements or propositions as states. I have accordingly taken the notion of a belief state to be those belief states that something or other is a so-and-so.

¹⁴ The extension of the present theory to finite sets of states that might involve several individuals is fairly straightforward.

¹⁵ In Mellor (1990, p. 40). There's another but similar story that could be told about the caterpillar. But the chicken came first.

if $tbtp$ is identical to $tbtq$, it will follow from our account of the truth and falsity of belief states, that p and q have to be equivalent. But the converse may be false. Our theory leaves that issue open, as we think it should.¹⁶

The assumptions in the Tarskian construction were not substantive. The corresponding assumptions for these belief states may be more controversial. The assumption of identities such as $tbtq = tbtp$ require that the term $tbtq$ refers to a mental state that exists. I also assume that although a person can be in many different belief states at one time, that there is only one belief state that p for a person at a time. As for the conditions of non-identity that are assumed, they too seem to come at some price. Parallel to the idea that if p and q are different, then so are their canonical names “ $|p|$ ” and “ $|q|$ ”, we assume that if two belief states are different, then the canonical terms that refer to those states are different – i.e. $tbtq \neq tbtp$.¹⁷

The theory so far, is silent about what the ps and qs may be. They could be sentences, statements, or propositions. We certainly don’t want to say that they have to be truth bearers. That would doom Ramsey’s project from the start. Ramsey aim was to give an account of truth and falsity of belief states such as the belief state that p ($tbtp$). To assume the notion of “truth bearer” would use the concept of truth for (say) propositions to elucidate the truth or falsity of certain belief states. That would run the analysis in precisely the wrong direction. Fortunately, for the limited purposes at hand, we do not have to resolve these issues.

We shall assume that we have a finite set of belief states $\Delta_n = \{tbtp_1, tbtp_2, \dots, tbtp_n\}$ for which it is assumed that $tbtq_i = tbtp_i$ (for $1 \leq i \leq n$), and $tbtq_i \neq tbtp_j$, for $p_i \neq p_j$ (for $1 \leq i, j \leq n$).

Let us begin with the simplest case – when the task is to define a truth predicate of belief states for a set of belief states that has only one member, say $tbtp$. Recall Ramsey’s stress on the centrality of the idea that “... a belief is true if it is a belief that p , and p ”. We have to make an adjustment of this “truism” about beliefs (his description) to take into account that now the problem is to define truth and falsity for belief states. The natural expression of the insight for belief states is then given by:

$$\text{Tr}(x) : (x = tbtp) \wedge p.$$

This tells us, when a set of belief states has only one member, say $tbtp$, then any

¹⁶ That is, there could be p and q of the same truth value, and only slightly different in what they say, but the corresponding belief states might be identical. It’s a kind of doxastic belief threshold phenomenon.

¹⁷ It would be interesting to determine whether these states of belief should be taken as tokens or types. I would prefer types since I think a person can be in the same state twice, and different people can be in the same state. I don’t think that the construction of a truth predicate for finite sets of states of beliefs forces the issue one way or the other.

belief state x is true just in case it is the belief state that p , and p . If the set of belief states Δ , has more than one member, then the truth predicate for that set of states is defined by the disjunction of clauses for the states of Δ , as if they were each considered as in the single case. More precisely, if $\Delta = \{tbtp, tbtq, \dots, tbtr\}$, then the truth predicate for the set of states Δ is defined this way:

$$\text{Tr}_{\Delta}(x) : [(x = tbtp) \wedge p] \vee [(x = tbtq) \wedge q] \vee \dots \vee [(x = tbtr) \wedge r],$$

where “ $\text{Tr}_{\Delta}(x)$ ” says that x is a true belief state of Δ .¹⁸

Earlier we noted the passage in *Facts and Propositions* in which Ramsey expressed his belief that the truth or falsity of the mental factor of a belief (judgement) depends only on what proposition it is that is judged. Ramsey states what that dependence is for an illustrative special case: the belief that aRb is true, if aRb , and false if not.

If we make an adjustment in this passage to reflect that it is the truth and falsity of belief states that has to be elucidated, then Ramsey’s observation is that the truth or falsity of the mental factor (e.g. belief state) depends only on what proposition it is that is believed. Transposed to belief states rather than beliefs, this would become: the belief state that aRb is true if aRb , and false if not. More generally (and schematically expressed), the belief state that p is true if and only if p . I shall call this *Ramsey’s Belief Schema*:

(RBS): For any belief state $tbtp$, $\text{Tr}(tbtp) \leftrightarrow p$.

I think that every instance of the Ramsey Belief-Schema is provable. Consider first the simplest case: the truth predicate for a set of belief states for which we seek a predicate “true” has only one member. In this very simple case, the Ramsey “truism” can be expressed as

$$(1) \quad \text{Tr}(x) : (x = tbtp) \wedge p.$$

In this case then, the argument for **(RBS)** is simple: To show that $\text{Tr}(tbtp) \rightarrow p$, suppose that $\text{Tr}(tbtp)$. Then by our definition, $(tbtp = tbtp) \wedge p$, and consequently, p . In short we have one half of the belief schema: It is surprising that the converse also holds: that is $p \rightarrow \text{Tr}(tbtp)$.

The proof goes this way: suppose that p . We have assumed that the $tbtps$ under discussion all refer to belief states of an individual. In that case we have the identity statement $(tbtp = tbtp)$. So we have $(tbtp = tbtp) \wedge p$. Consequently, $\text{Tr}(tbtp)$. Thus in the single case, we have proved that $\text{Tr}(tbtp) \leftrightarrow p$.

¹⁸ “is a true belief state” is ambiguous, between “is a true (belief state)” and is a (true belief) state”. We mean the former of course - we want to pursue Ramsey’s idea of ascribing truth or falsity to belief states.

The idea of defining a truth predicate for belief states, one-at-a-time, proves unsatisfactory for the very same reasons that defining a truth predicate for sentences, one at a time is unsatisfactory. The obvious thing to do is to abandon the one-at-a-time procedure, and define the truth predicate for arbitrarily large finite sets of belief states.

Let $\Delta_n = \{tbt p_1, tbt p_2, \dots, tbt p_n\}$ be a finite set of belief states. The truth predicate for Δ_n (suppressing the notation for the set), is

$$(2) \quad \text{Tr}(x): [(x = tbt p_1) \wedge p_1] \vee [(x = tbt p_2) \wedge p_2] \vee \dots \vee [(x = tbt p_n) \wedge p_n].$$

In this case for each of the p_i in $\{p_1, p_2, \dots, p_n\}$, we can prove that $\text{Tr}(tbt p_i) \leftrightarrow p_i$.

The proof is again straightforward. There are only n members of Δ_n , and it is assumed that all the $tbt p_i$ s are distinct from each other. Then for any p_k ,

$$\text{Tr}(tbt p_k): [(tbt p_k = tbt p_1) \wedge p_1] \vee [(tbt p_k = tbt p_2) \wedge p_2] \vee \dots \vee [(tbt p_k = tbt p_n) \wedge p_n].$$

Since the negation of every disjunct other than $[(tbt p_k = tbt p_k) \wedge p_k]$, is provable, we conclude that $\text{Tr}(tbt p_k) \rightarrow [(tbt p_k = tbt p_k) \wedge p_k]$, is provable, and consequently, so too is $\text{Tr}(tbt p_k) \rightarrow p_k$. Conversely, assume that p_k . Since we have $tbt p_k = tbt p_k$, we have the conjunction $(tbt p_k = tbt p_k) \wedge p_k$. But this conjunction is a disjunct of $\text{Tr}(tbt p_k)$, and so we have $\text{Tr}(tbt p_k)$. That is, we have proved that $p_k \rightarrow \text{Tr}(tbt p_k)$. Consequently, $\text{Tr}(tbt p_k) \leftrightarrow p_k$.

In light of this result, the more exact description of (RBS) should be this:

(RBS)*: Let $\Delta_n = \{tbt p_1, tbt p_2, \dots, tbt p_n\}$ be a finite set of belief states of an individual, and “Tr” be the truth predicate for Δ_n , then for any state $tbt p_i$ in Δ_n , $\text{Tr}(tbt p_i) \leftrightarrow p_i$ is provable.

The definition of the truth predicate for sets of belief states does not have any occurrences of “truth” in it, for the very same reason that Tarski’s definition of truth doesn’t. It therefore satisfies one of the desiderata of Ramsey’s program for connecting belief states and truth.

The Ramsey belief schema for belief states is striking, and it lays to rest a possible suspicion that the shift to true belief states is just another way of referring to true beliefs. That is, suppose that $tbt p$ (the belief state that p) is a state of belief of Oscar. One could suppose that that state is true if and only if Oscar believes that p . That is,

$$\text{Tr}(tbt p) \text{ is equivalent to “Oscar believes that } p\text{.”}$$

However this equivalence of the truth of Oscar's state of belief and Oscar's belief is incompatible with the Ramsey Belief Schema (RBS)*. "Tr(tbtp)" implies p, but "Oscar believes that p" doesn't. If Oscar's belief state that p is true, then it follows (by the Ramsey Belief Schema) that p, but p doesn't follow simply from the statement that Oscar believes it. This seems to me to be correct. If Oscar believes that p, then Oscar may very well be in the belief state that p. Even if he is, even if he believes that p is true, it doesn't follow that that belief state is true, nor that p is true. The shift to states of belief and their truth seems to yield more than simply adhering to the truth of belief statements.

5.

Some Possible Reservations. In the course of proving the various instances of the Ramsey Belief Schema we made several assumptions that might be thought to be incompatible with Ramsey's overall views, or might be thought to be just plain wrong. Here are several of interest.

(1) **Finitism.** It is worth noting that although the restriction of truth predicates to finite sets of sentences might seem too stringent to all but someone who denied the existence of any actual infinite, the restriction to finite sets of belief states may be acceptable. After all, the set of belief states can be as large as one wants or needs, and it is a bit of a stretch to insist that there is a real need to allow for beings with infinitely many beliefs states. Furthermore, there is some reason to believe that Ramsey in 1929 endorsed a finitist view which rejected the existence of any actual infinite.¹⁹ Thus the restriction to finite sets would be entirely compatible with Ramsey's general philosophical commitments held in 1929.

(2) **One Truth Predicate of Sets of Belief States, or Many.** When the notion of truth is elucidated for arbitrarily large finite sets of belief states, it becomes clear that the truth predicates for different sets them will be different. This is also the case when truth predicates are elucidated for various finite sets of sentences. However in both cases the notion of truth is *cummulative*. If one set of belief states is a subset of a second, then all those belief states which are true with respect to the truth predicate of the first set will also be true with respect to the truth predicate for the second. That should ameliorate the systematic ambiguity that is involved.

¹⁹ As reported by R. B. Braithwaite (1931, p. xii).

(3) **Conjunctive Closure and Negation Completeness.** According to another possible objection, the Ramsey Belief Schema has consequences that are controversial and may not be correct. Two examples:

$$(i) \quad (\text{Conjunction}) \quad \text{Tr}(\text{tbtp}) \wedge \text{Tr}(\text{tbtq}) \leftrightarrow \text{Tr}(\text{tbt}[p \wedge q]),$$

and

$$(ii) \quad (\text{Negation}) \quad \text{Tr}(\text{tbtp}) \leftrightarrow \neg \text{Tr}(\text{tbt}[\neg p]).$$

Consider (i) first. This condition on belief states looks like the familiar controversial conjunctive condition on beliefs: a person believes a conjunction if and only if they believe each of the conjuncts. However, this result is quite different. One has to be careful and not confuse the usual conjunctive rule for beliefs with the present result about the way the truths of the states tbtp , tbtq , and $\text{tbt}[p \wedge q]$ are related. A proof of (i) might run this way: By **(RBS*)**, we have $\text{Tr}(\text{tbtp}) \leftrightarrow p$, $\text{Tr}(\text{tbtq}) \leftrightarrow q$, and $\text{Tr}(\text{tbt}[p \wedge q]) \leftrightarrow p \wedge q$. Therefore by classical sentential logic, we have $\text{Tr}(\text{tbtp}) \wedge \text{Tr}(\text{tbtq}) \leftrightarrow \text{Tr}(\text{tbt}[p \wedge q])$.

The proof is not faulty, but one has to pay careful attention to what it assumes. It is assumed that all three states of belief, tbtp , tbtq , and $\text{tbt}[p \wedge q]$, are in some set Δ , and “Tr” is the truth predicate for Δ . The result holds if all three states are in the set. However, if tbtp and tbtq are in some set, but the belief state $\text{tbt}[p \wedge q]$ is not, then (i) will not be provable.

In other words, it is not guaranteed by this theory that if someone is in the belief state tbtp , and also in the state tbtq , then they are also in the conjunctive state $\text{tbt}[p \wedge q]$. The theory so far leaves that open –as it should.

Condition (ii) on negation also appears to be controversial. In the case when the belief states are of some particular individual, it seems to imply that for any p , that the individual is either in the belief state that p , or else in the belief state that $\neg p$. One would have to be in either one belief state or the other. There’s no room for a kind of indeterminacy, and that would be a drawback of the theory. However the present result is quite different. It says that of the two states of some individual, tbtp , and $\text{tbt}[\neg p]$, one or the other of them is true: $\text{Tr}(\text{tbtp}) \vee \text{Tr}(\text{tbt}[\neg p])$.

The proof of (ii) is simple enough. Suppose that Δ is some set of belief states, and that for some p , both tbtp and $\text{tbt}[\neg p]$ are in Δ . Then if “Tr” is the truth predicate for Δ , we have both $\text{Tr}(\text{tbtp}) \leftrightarrow p$, and $\text{Tr}(\text{tbt}[\neg p]) \leftrightarrow \neg p$. Consequently $\text{Tr}(\text{tbtp}) \leftrightarrow \neg \text{Tr}(\text{tbt}[\neg p])$.²⁰ The proof is fine, but it does depend

²⁰ The assumptions under which (ii) and the claim that $\text{Tr}(\text{tbtp}) \vee \text{Tr}(\text{tbt}[\neg p])$ are provable also show that they are equivalent to $p \leftrightarrow \neg \neg p$, and $p \vee \neg p$ respectively. That isn’t a problem, unless intuitionist scruples are at issue.

upon both states, $tbtp$ and $tbt[\neg p]$ being members of the set for which the truth predicate is defined. If one of them is in the set but the other not, then (ii) will not be provable.

(4) **Incoherent States of Belief.** Some of the preceding observations may suggest some further misgivings about the theory we have been explaining. We have considered various sets of belief states and the truth predicates for them. Some of those sets might seem to be strange, perhaps incoherent sets of belief states, if they are supposed to be states of some one individual. The suggestion is that there needs to be further work to delimit or restrict the membership of collections of states, before we define truth predicates for them. For example, in the discussion of (ii) above, we considered certain sets of belief states that contained the two states of belief $tbtp$ (the belief state that p) and $tbt[\neg p]$ (the belief state that $\neg p$). Could they both be states of belief of some individual? The theory developed thus far imposes no constraints on the belief states that might be collected into a set, and have a truth predicate for them.²¹ Thus, for all that we have said so far, $tbtp$ and $tbt[\neg p]$ could be states of one individual. That has not been ruled out. What has been ruled out, however, is that those two states cannot both be true states. The reason is simply that the conjunction of $Tr(tbtp)$ and $Tr(tbt[\neg p])$ is inconsistent since it implies $p \wedge \neg p$. One of the two states is not true – say it is $tbt[\neg p]$. Although it is a false state of belief, it is nevertheless still a state of belief of the individual. It may be the understatement of the year, but not all states of belief of an individual have to be true.

(5) Two Kinds of Truth Predicates (for “Sentences” and for States of Belief) or One?

There is a temptation to think that despite our emphasis on the advantages of shifting to states of belief rather than belief sentences, there isn't much difference or advantage in defining truth predicates for finite sets of belief states. If that were so, then the point of trying to give an account of the truth of belief states that did not rely on an account of truth for sentences, statements, or even propositions, would be pointless. The reason involves a mistaken inference from the two kinds of schemata. Assume that $\Delta = \{p, q, \dots, r\}$ is some finite set of “sentences”, and that $\Delta^* = \{tbtp, tbtq, \dots, tbtr\}$ is the corresponding finite set of states of belief. This is just an assumption. We do not assume that for every set of type Δ , there will always exist a corresponding Δ^* .

If we form the corresponding truth predicates for the two sets, we will have proofs of the following instances of each schema:

²¹ In this way, the treatment for finite sets of states of belief and finite sets of sentences are on a par.

(i) $\text{Tr}^*(\text{tbt}p) \leftrightarrow p$,

and

(ii) $\text{Tr}(|p|) \leftrightarrow p$.

From the two schemata we could infer the equivalence

(iii) For the sets Δ and Δ^* , $\text{Tr}^*(\text{tbt}p) \leftrightarrow \text{Tr}(p)$.

This however does not entail that the two truth predicates are the same; they are not even coextensional for one is defined on belief states, and the other is defined on “sentences”.

There may be a deeper problem with (iii) that has to do with additional assumptions about belief states other than those we made in deriving the equivalence. Here is a made-up example of what we have in mind. Suppose that someone has a theory of belief states according to which belief states are maximal in this sense: For any disjunction, the belief state that $(A \vee B)$ is either identical to the belief state that A or it is identical to the belief state that B (i.e. $\text{tbt}[A \vee B] = \text{tbt}A$, or $\text{tbt}[A \vee B] = \text{tbt}B$). We might call such a view “dedicated belief state intuitionism”

We can now see that (iii) will fail for certain disjunctions: Let the disjunction $(A \vee B)$ and its disjuncts (neither of which implies the other), be in the set Δ , and the belief state $\text{tbt}[A \vee B]$ as well as the belief state that A , and the belief state that B , be in the corresponding set Δ^* of belief states. Suppose too that (iii),

$$\text{Tr}(A \vee B) \leftrightarrow \text{Tr}^*[\text{tbt}(A \vee B)],$$

where “Tr” is the truth predicate defined for the set of sentences Δ , and “Tr*” is the truth predicate defined for the set of corresponding belief states Δ^* . “Tr($A \vee B$)” is equivalent of course to “($A \vee B$)”. Consider $\text{Tr}^*[\text{tbt}(A \vee B)]$. Either $\text{tbt}(A \vee B)$ is identical to $\text{tbt}A$ (the first case) or it is identical to $\text{tbt}B$ (the second case). In the first case, $\text{Tr}^*[\text{tbt}(A \vee B)]$ is equivalent to $\text{Tr}^*[\text{tbt}A]$, which is equivalent to A . Consequently $(A \vee B)$ implies A , and so B implies A . But that is impossible. In the second case, $\text{Tr}^*[\text{tbt}(A \vee B)]$ is equivalent to $\text{Tr}^*[\text{tbt}B]$ which is equivalent to B , so that $(A \vee B)$ implies B . Consequently A implies B . But that too is impossible. Therefore (iii) fails, and it fails because in this imagined example of a theory of belief states, there are no disjunctive states of belief where neither disjunct implies the other.

The moral of this imaginary extension of our simple theory of belief states is that one has to be careful in how a theory of belief states and their truth are developed beyond the elementary assumptions in our account. This is especially

so, if the theory of belief states says, as our made up example does, that certain kinds of belief states do not exist.

6.

The Unfinished Theory and Pragmatism. The theory as presented thus far, yields certain results that were characteristic of Ramsey's thoughts about the relation of belief states and truth. By focusing on the elucidation of an account of the truth and falsity of mental states we were able to construct arguments for certain theses like the Ramsey Belief Schema* which Ramsey advocated. However, in our proof that the instances of the Ramsey Schema were provable, we used certain assumptions that in turn need some explanation. We assumed the existence of finite sets of belief states $tbtp$ etc., since we assumed that they were self identical ($tbtp = tbtp$, etc.), and they were different from each other ($tbtp \neq tbtpq$, etc.). Just to fix our ideas a little more definitely, consider the truth predicate for a set of two belief states $tbtp$ and $tbtpq$ of an individual. That truth predicate was given by $Tr(x) : [(x = tbtp) \wedge p] \vee [(x = tbtpq) \wedge q]$. Consequently if we want to determine whether $Tr(tbtp) \leftrightarrow p$, and $Tr(tbtpq) \leftrightarrow q$ hold or not, we need information as to whether $tbtp = tbtp$, $tbtpq = tbtpq$, and $tbtp \neq tbtpq$. How are we supposed to obtain that information? Something should be said about when we attribute a belief state to an agent, and when the beliefs states that we attribute are different.

Ramsey had already addressed a similar problem in *Facts and Propositions*. The answer he thought lay in a so-called pragmatist view that was part of his story of the celebrated chicken and the caterpillar. From that story and a hint at the kind of pragmatism he had in mind, we can see one likely way in which the ascription of beliefs to individuals, no matter where they are in the pecking order, can be carried over with some adjustment, to the case of the attribution of belief states to them.

In order to focus on what role Ramsey assigned to the "pragmatist view", it is best to begin with two passages from "*Facts and Propositions*". The first is the celebrated passage on that famous chicken, and the second implicates Russell as the source of his pragmatism. Each is worth quoting in full:

"... It is, for instance, possible to say that a chicken believes a certain sort of caterpillar to be poisonous, and mean by that merely that it abstains from eating such caterpillars on account of unpleasant experiences connected with them. The mental factors in such a belief would be parts of the chicken's behaviour, which are somehow related to the objective factors, viz. the kind of caterpillar and poisonousness. An exact analysis of this relation would be very difficult. But it might well be held that in regard to this kind of belief the pragmatist view was correct, i.e. that the

relation between the actions were such as to be useful if, and only if, the caterpillars were actually poisonous. Thus any set of actions for whose utility p is a necessary and sufficient condition might be called a belief that p , and so would be true if p , i.e. if they are useful.”²²

Three things are worth noting about the first passage. It attributes a definite belief to the chicken: a certain sort of caterpillar is poisonous. Second, there is mention of the mental factor and the objective factors of that belief. We take the mental factor to be a reference to what he later called the belief state (tbtp), and the objective factors concern the kind of caterpillar and its poisonousness (p). Third there is the idea that a pragmatic view will provide an account of why that particular belief is properly attributed to the chicken. In the second passage, Ramsey is explicit about the Russellian origin of his pragmatism:

“My pragmatism is derived from Mr. Russell; and is of course, very vague and undeveloped. The essence of pragmatism I take to be this, that the meaning of a sentence is to be defined by reference to the actions to which asserting it would lead, or, more vaguely still, by its possible causes and effects. Of this I feel certain, but of nothing more definite.”²³

Of course we cannot discount Russell’s influence, but as we shall see, Ramsey’s view seems closer to a Peircean brand of pragmatism.²⁴ It is worth a try to include caterpillars as having beliefs, but the Russellian suggestion to define the meaning of sentences by reference to actions that lead to their assertions is off the mark. However, the relevant actions needn’t be limited to actions that lead to assertions, and Ramsey considers a more inclusive kind of action. He says that what we mean by attributing to the chicken a belief that those caterpillars are poisonous is that it avoids eating them on account of the untoward experiences that eating them would provide. The reference to certain actions of the chicken provides a way of attributing specific beliefs to the chicken. It is intended as a solution to the problem of belief ascription.

Even if this so-called pragmatic appeal to actions indicates whether the individual has that belief, it does not of course settle the question of whether that belief is true. It might be true that the chicken believed the caterpillar was poisonous, and just be wrong. Lucky caterpillar!

Moreover, the project is, as Ramsey said in his draft of “The Nature of Truth”, to elucidate the truth (or falsity) of belief *states*. Consequently prag-

²² Mellor (1990, p. 40).

²³ Mellor (1900, p. 51).

²⁴ The case for the Peircean connection is clearly explained in Sahlin (1990, p. 70-73), together with an interesting relation of it to decision theory. This view of the matter is also supported by similar considerations in Dokic and Engel (2001, p. 22-25).

matist accounts of either the meaning or the truth of sentences, statements, or propositions would seem to be irrelevant. It is belief *states* whose truth needs discussion.²⁵ There is, nevertheless, a very appropriate place in this account, where Peircean views help to move things along, by saying something about the attribution of belief states, and the way in which one belief state might be distinguished from another. The idea is to continue to deploy a pragmatist view even if we shift from a discussion of sentences such as “Oscar believes that caterpillars are poisonous” to the attribution of a belief state (the belief state that caterpillars are poisonous) to Oscar.

There are some things that Peirce wrote, in “How to make our Ideas Clear”²⁶ that are helpful. One idea is that

“The essence of belief is the establishment of a habit, and different beliefs are to be distinguished by the different modes of action to which they give rise” (pp. 129–130),

and

“... what a thing means is simply what habits it involves. Now, the identity of a habit depends on how it might lead us to act, not merely under such circumstances as are likely to arise, but under such as might possibly occur, no matter how improbable they may be. ... Thus we come down to what is practical and tangible, as the root of every real distinction of thought, no matter how subtle it may be; and there is no distinction of meaning so fine as to consist in anything but a possible difference of practice.” (p. 131)

The idea is then to transpose these suggestions to belief states, and propose that the states of belief (of Oscar, say) are connected with certain dispositions of Oscar to act under various circumstances, even improbable ones. Being specific about the kind of connection is of course the big problem. Rather than revert

²⁵ The issue of whether truth is a byproduct of an account of beliefs or the byproduct of an account of psychological mental states like belief states is not easily separated by citing those Ramsey papers in which beliefs are the target, or papers in which beliefs states are the target. In “Facts and Propositions” (1927), Mellor (1990), in the chicken passage, there is mention of both beliefs and their mental factors, and in the draft of “The Nature of Truth” (1927 – 28, or perhaps 29), Rescher and Majer (1991), the discussion is clearly intended to be about mental states, though there too he also talks about all beliefs as necessarily having a *propositional reference* (“that they are all beliefs that something is so-and-so”, p. 7). It is possible that he could be also be understood, as also saying that all belief states are belief states that something is so and so, which is the way we understand him here.

²⁶ Hauser and Kloesel (1992).

to meanings, for the present let us just assume a weak Peirce-like proposal: the actions under various circumstances are *evidence* that Oscar is in a certain state of belief (that the caterpillar is poisonous).²⁷ So appeal to evidence of a practical, tangible sort would be evidence that the chicken was in a certain belief state. And evidence that one state of belief is different from another would be provided by the possible difference of practice associated with each. At root the difference, Peirce says, comes down to what is practical and tangible. That emphasis on the practical and tangible would indicate the utility these belief states have for the chicken, and us, and might be what lay behind Ramsey's use of the notions of utility and usefulness at the end of the caterpillar passage. The nice feature of Ramsey's appeal to pragmatic considerations is that it allows us to secure evidence for the claim that a certain state of belief is true, i.e. Tr(tbtp), without relying on some account of truth already in place. To my mind the present theory is coherent. With a little luck, it may even be the beginning of a theory that is correct.

Arnold Koslow
The Graduate Center
The City University of New York
365 Fifth Avenue, New York
akoslow@mindspring.com

References

- Braithwaite, R. B., (Ed.) (1931). *The Foundations of Mathematics and other Logical Essays*, by F. P. Ramsey. London: Routledge.
- Dokic, H., & Engel, P. (2001). *Frank Ramsey, Truth and Success*. London: Routledge.
- Hause, N., & Kloesel, C. (Eds.) (1992). *The Essential Peirce. Selected Philosophical Writings* (vol. I, 1867–1893). Bloomington: Indiana University Press.
- Koslow, A. (1990). Ramsey on Truth and Simplicity. In Nils-Eric Sahlin (Ed.). *Ramsey's Ontology. Metaphysica*, 3, 89–108.

²⁷ The usual connection by pragmatists is in terms of meanings. However we are concerned with belief states, and their connection to various acts is not a matter of meaning. Ramsey's appeal to causes would lose the connection with what is tangible and practical –i.e. to utility. The use of an evidential connection however, has its drawbacks as well.

- Mellor, D. H. (Ed.) (1990). *F.P. Ramsey, Philosophical Papers*. Cambridge University Press.
- Majer, U. (1991). Ramsey's Theory of Truth and the Truth of Theories: A Synthesis of Pragmatism and Intuitionism in Ramsey's Last Philosophy. *Theoria*, 57(3), 162-195.
- Rescher, N., & Majer, U. (Eds.) (1991). On Truth. *Original Manuscript Materials (1927-1929) from the Ramsey Collection at the University of Pittsburgh*. Kluwer Academic Publishers.
- Sahlin, N.-E. (1990). *The Philosophy of F.P. Ramsey*. Cambridge University Press.
- Sahlin, N.-E. (Ed.) (2005). Ramsey's Ontology. *Metaphysica*, 3.
- Tarski, A. (1969). Truth and Proof. *Scientific American*, 6, 63-77. The revised version can be found in S. R. Givant and R. N. McKenzie (1986).
- Givant, S. R., & McKenzie, R. N. (Eds.) (1986). *Tarski's Collected Works*. Basel: Birkhäuser.
- Hause, N., & Kloesel, C. (Eds.) (1992). The Essential Peirce. *Selected Philosophical Writings* (vol. I, 1867-1893). Bloomington: Indiana University Press.