

# ON THE CONSISTENCY OF CHISHOLM'S AXIOMS OF EPISTEMIC LOGIC

DAN HICKS

In the opening chapters of (Chisholm, 1977) and (Chisholm, 1982), Chisholm sketches an account of some basic concepts of epistemic justification in terms of the relation  $\ulcorner \varphi$  is more reasonable than  $\psi \urcorner$  (suppressing references to epistemic agents and moments of time since they are not relevant for my purposes here), where  $\varphi$  and  $\psi$  are (roughly) propositional attitudes logically built from the primitives  $\ulcorner$ believing  $p\urcorner$  and  $\ulcorner$ withholding  $p\urcorner$ , for any proposition  $p$ . In particular, Chisholm gives what he calls axioms for the logic of this relation. In this paper I examine the consistency of these axioms.

## 1. CEP AND ITS RELATIVES

Let  $\mathcal{L}'$  be any formal language.  $\mathcal{L}'$  provides the sentences (or, if one prefers, propositions) towards which the epistemic agent takes her or his doxastic attitudes. The relative reasonability of these attitudes is expressed using the formal language  $\mathcal{L}$ . The relationship between  $\mathcal{L}'$  and  $\mathcal{L}$  is the relation of object language to meta-language.

$\mathcal{L}$  is a first-order formal language. The variables of  $\mathcal{L}$  range over the sentences of  $\mathcal{L}'$ , and are written using lower-case Italic letters  $p$ ,  $q$ , &c.  $\mathcal{L}$  contains, besides the ordinary first-order logical operators, two unary doxastic operators  $b$  and  $w$  – expressing ‘believing’ and ‘withholding’, respectively – and two binary operators  $>$  and  $=$  – expressing ‘is more reasonable than’ and ‘is just as reasonable as’.

In general, we will consider sentences of  $\mathcal{L}$  where the variables only occur within the scope of the unary operators, and the unary operators only occur within the scope of the binary operators. That is, we will only be considering sentences built out of comparisons between doxastic attitudes using the logical connectives, eg,  $(b(p) > b(q)) \vee (w(p) > b(q))$ .

For convenience, we introduce the additional operators  $d$  and  $\leq$ .

$$d(p) =_{\text{df}} b(\neg p) \tag{1.1}$$

$$\varphi \leq \psi =_{\text{df}} \neg(\varphi > \psi) =_{\text{df}} (\varphi = \psi) \vee (\psi > \varphi), \tag{1.2}$$

where  $\varphi$  and  $\psi$  are schematic variables for doxastic attitudes. Note that we do *not* formally include doxastic attitude variables in  $\mathcal{L}$ . Their use here is simply for convenience of presentation. The dual definition of  $\leq$  is justified by the order axioms (A1) and (A2): it is clear that Chisholm intends  $>$  to be a linear order.

**1.1. CEP.** We start by transcribing Chisholm and Keim's formal system CEP ('Calculus of Epistemic Preferability'<sup>1</sup>) into the language  $\mathcal{L}$ , taking advantage of the opportunity to fix some typographical errors. CEP is the theory in  $\mathcal{L}$  with the following seven axioms (referring to the schema (A1) and (A2) as axioms for the sake of simplicity):

$$(A1) \quad \forall \varphi, \psi ((\varphi > \psi) \rightarrow \neg(\psi > \varphi))$$

The relation  $>$  is asymmetric.

$$(A2) \quad \forall \varphi, \psi, \xi (((\varphi \leq \psi) \wedge (\psi \leq \xi)) \rightarrow (\varphi \leq \xi))$$

The relation  $>$  is transitive.

$$(A3) \quad \forall p, q ((b(p) > b(q)) \leftrightarrow (d(q) > d(p)))$$

If believing  $p$  is more reasonable than believing  $q$ , then disbelieving  $q$  is more reasonable than disbelieving  $p$ .

$$(A4) \quad \forall p ((w(p) \leq b(p)) \rightarrow (b(p) > d(p)))$$

If remaining agnostic with respect to  $p$  is no more reasonable than believing  $p$ , then believing  $p$  is more reasonable than disbelieving  $p$ .

$$(A5) \quad \forall p, q ((w(p) = w(q)) \leftrightarrow ((b(p) > b(q)) \vee (d(p) > b(q))))$$

Remaining agnostic with respect to  $p$  is just as reasonable as remaining agnostic with respect to  $q$  if, and only if, either believing  $p$  is more reasonable than believing  $q$  or disbelieving  $p$  is more reasonable than believing  $q$ .

$$(A6) \quad \forall p, q (((b(p) > b(q)) \wedge (b(p) > d(q))) \rightarrow (w(q) > w(p)))$$

If believing  $p$  is more reasonable than both believing and disbelieving  $q$ ,

---

<sup>1</sup>Chisholm & Keim, 1972, 101

then remaining agnostic with respect to  $q$  is more reasonable than remaining agnostic with respect to  $p$ .

$$(A7) \quad \forall p (w(p) = w(\neg p))$$

Remaining agnostic with respect to  $p$  is just as reasonable as remaining agnostic with respect to  $\neg p$ .

There are several other formal systems which we will consider.

1.2. **CEP<sup>-</sup>**. CEP<sup>-</sup> is the formal system with axioms (A1)-(A3) and (A5)-(A7), ie, it is CEP without axiom (A4).

1.3. **GT**. GT is the extension of CEP<sup>-</sup> with the addition of the following axiom:

$$(\forall T) \quad \forall p (b(p) > w(p))$$

For all  $p$ , believing  $p$  is more reasonable than remaining agnostic with respect to  $p$ .

That is, it is CEP where axiom ( $\forall T$ ) replaces (A4).

1.4. **CEP<sup>+</sup>**. CEP<sup>+</sup> is the extension of CEP<sup>-</sup> with the addition of the following two axioms:

$$(\exists T) \quad \exists p ((b(p) > w(p)) \wedge (d(p) > w(p)))$$

There is at least one  $p$  such that believing  $p$  is more reasonable than remaining agnostic with respect to  $p$  and disbelieving  $p$  is more reasonable than remaining agnostic with respect to  $p$ .

$$(\exists E) \quad \exists p ((w(p) \leq b(p)) \rightarrow (b(p) > d(p)))$$

There is at least one  $p$  such that, if remaining agnostic with respect to it is no more reasonable than believing it, then believing it is more reasonable than disbelieving it.

1.5. **Chisholm's terms of epistemic appraisal**. Finally, in both (Chisholm & Keim, 1972) and (Chisholm, 1977), Chisholm gives a list of definitions for terms of 'epistemic appraisal'. We shall transcribe his definitions into  $\mathcal{L}$ .

**beyond reasonable doubt:**  $R(p) =_{df} b(p) > w(p)$

$p$  is beyond reasonable doubt, by definition, if believing  $p$  is more reasonable

than remaining agnostic with respect to  $p$ . Other definitions are transcribed into English similarly.

**acceptable:**  $A(p) =_{\text{df}} w(p) \leq b(p)$

**gratuitous:**  $G(p) =_{\text{df}} b(p) \leq w(p)$

**unacceptable:**  $U(p) =_{\text{df}} \neg A(p)$

**ought to be withheld:**  $W(p) =_{\text{df}} (w(p) > b(p)) \wedge (w(p) > d(p))$

**counterbalanced:**  $C(p) =_{\text{df}} b(p) = d(p)$

**some presumption in its favour:**  $F(p) =_{\text{df}} b(p) > d(p)$

## 2. (A4)

We begin our substantive discussion by examining axiom (A4). Suppose that  $w(p) \leq b(p)$  and  $w(\neg p) \leq b(\neg p)$  for at least one  $p$ . That is, remaining agnostic with respect with  $p$  is no more reasonable than each of believing  $p$  and disbelieving  $p$ . Then, by one application of (A4), we have

$$b(p) > d(p). \tag{2.1}$$

Next, by a second application, we have

$$b(\neg p) > d(\neg p). \tag{2.2}$$

But, by (2), (8) is

$$d(p) > b(p). \tag{2.3}$$

That is,

$$(b(p) > d(p)) \wedge (d(p) > b(p)), \tag{2.4}$$

contradicting the asymmetry axiom (A1).

For example, suppose that relative reasonability is measured by reference to a scale of absolute reasonability, where 0 is minimally reasonable ('no attitude could ever be less reasonable') and 10 is maximally reasonable ('no attitude could ever

be more reasonable'). Let  $p$  be the proposition that  $1 + 1 \neq 2$ . Then we make the following absolute reasonability assignments:

$$\begin{array}{ll} b(p) \mapsto 0 & b(\neg p) \mapsto 10 \\ d(p) \mapsto 10 & d(\neg p) \mapsto 0 \\ w(p) \mapsto 0 & w(\neg p) \mapsto 0 \end{array}$$

That is, denying  $1 + 1 \neq 2$  is maximally reasonable (in Chisholm's terminology, this denial is absolutely certain), and affirming it or remaining agnostic about it are as unreasonable as could be. And the antecedent in (A4) is satisfied:  $w(p) \leq b(p)$ , since  $0 \leq 0$ . But the consequent is false, indeed as false as could be:  $b(p) \not\leq d(p)$ , believing  $p$  is very much not more reasonable than denying  $p$ . Denying  $p$  is the only option on the table.

Of course, we didn't need the numerical reasonability assignments here. Instead, we could have simply claimed that withholding judgement with respect to  $1 + 1 \neq 2$  is no more reasonable than believing that  $1 + 1 \neq 2$  (they're both completely preposterous, either equally or such that withholding judgement is worse than believing), and withholding judgement with respect to  $1 + 1 \neq 2$  is no more reasonable than (indeed, is very much less reasonable than) believing that  $1 + 1 \neq 2$ .

Now, there are certainly epistemologically possible situations<sup>2</sup> in which withholding judgement about  $1 + 1 \neq 2$  is not the least reasonable option. If I am an anti-realist about mathematical entities, and affirming  $1 + 1 = 2$  would commit me to the existence of numbers, then withholding judgement with respect to  $1 + 1 \neq 2$

<sup>2</sup>The modality here is *not* what is normally called epistemic possibility, 'possible as far as  $X$  knows'. Instead, I intend epistemological possibility to be a species of alethic possibility: using the industry-standard possible worlds, the epistemologically possible worlds are a subset (or perhaps subclass) of the logically possible worlds, just as the nomologically possible worlds are a subset of the logically possible worlds. Thus, to say that it is epistemologically possible for  $X$  to be in circumstances  $C$  is to say that, among this subset of the logically possible worlds is one at which  $X$  is in circumstances  $C$ . Furthermore, this subset of the logically possible worlds is not determined relative to the knowledge of some agent, whether  $X$  or us who are considering  $X$  ('as far as  $X$  knows, she could be in any of these possible worlds'). Rather, this subset is the collection of just the logically possible worlds at which the axioms of epistemic logic are true. Again, this works like the nomologically possible worlds, which is the collection of just the logically possible worlds at which the laws of nature are true. One important corollary of this setup is that a proposed truth of epistemic logic is, if true, epistemologically necessarily true (true at all of the epistemologically possible worlds). Each of the counterexamples of this section should therefore be read as a claim that (A4) is not epistemologically necessarily true, and hence not true.

could very well be more reasonable than either affirming or denying it.<sup>3</sup> Then (A4) still holds.

However, the epistemological possibility of this situation (and I take it to be quite epistemologically possible) doesn't make the contrasting situation described in my counterexample epistemologically impossible. And the epistemological possibility of my counterexample is all I need for it to be a genuine counterexample, since (A4) asserts that such a situation is epistemologically impossible.

Here's a second counterexample. Suppose I pass a person who is obviously homeless on a very cold day, when the temperature is supposed to drop below zero that night. I'm wearing an old but still very, very warm coat, have a warm apartment to spend the night, and I was planning on buying a new coat tomorrow anyway. Do I have an obligation to give my coat to the homeless person? Say the answer to this question depends on whether I believe or disbelieve that the homeless person is a zombie (in David Chalmers' sense). Let  $p$  be the proposition that this being is a zombie.

I would probably be very reasonable in disbelieving  $p$ , and we might also suppose that I'm sceptical enough to also have some reasons for believing  $p$ . But let us add the assumption that my acting in this case (either giving the coat to the homeless person or passing them by) requires me to either affirm or deny  $p$ . More realistically, I only need the probabilistic belief that  $p$  is much less likely to be true than  $\neg p$ , or vice versa, but any analysis of this situation in the language  $\mathcal{L}$  (including an analysis by CEP) is going to require a simplifying assumption of this sort. Or perhaps the predicate ' $b$ ' could be redefined as the predicate 'is likely to be true with probability greater than  $p_0$ ' for some specific probability threshold  $p_0$ , though this could be incompatible with some of our intuitions about inferential belief, as in the lottery paradox.<sup>4</sup> In any case, I claim that it is epistemically possible to assume that I cannot remain agnostic with respect to  $p$ . I must make a doxastic choice, and then follow through on the consequences of that choice. Hence the ordering of reasonability is  $d(p) > b(p) > w(p)$ , contradicting (A4).

---

<sup>3</sup>I'd like to thank Alvin Plantinga for suggesting this case.

<sup>4</sup>Note that none of the formal systems discussed here have much to say about inferential belief.

Doing a search of the literature, I was surprised to discover that I was not the first to object to (A4) on the basis of these sorts of counterexamples. Both (Imlay, 1969) and (Ulm, 1975) make similar objections. Imlay considers the proposition 'I am in excruciating pain', when I am not actually experiencing pain. In this case, both remaining agnostic and affirming the proposition are as unreasonable as could possibly be; the only reasonable option is disbelieving it. Ulm's counterexample is 'I do not exist'. As Descartes would be the first to remind Chisholm, disbelieving this proposition is the only reasonable option.

In each case, the counterexample proposition is said to be such that both remaining agnostic and believing it are of 'lowest'<sup>5</sup> or 'minimal'<sup>6</sup> epistemic reasonableness. This is not quite necessary; all that is needed is for agnosticism to be no more reasonable than the other two doxastic attitudes. Then, for either  $p$  or  $\neg p$ , disbelieving will be more reasonable than believing, and (A4) will be false.

But even this is not quite necessary. (A4) asserts that situations like those in the counterexamples, situations where withholding judgement is the least reasonable doxastic option, are epistemologically impossible. Hence, (A4) fails not just if these situations actually obtain, but even if they are merely, and strictly, epistemologically possible. Thus it is no defence of (A4) to claim that, in other epistemologically possible situations, a different ordering of reasonability obtains; the given situations are epistemologically possible, and hence contradict (A4).

There are two conclusions to which these counterexamples might lead us. First, they show that CEP is logically inconsistent. We shall consider this possibility in the next section. Second, they show by *reductio ad absurdum* that CEP entails that remaining agnostic is never the worst option for any proposition, ie,

$$\text{CEP} \vdash \forall p ((w(p) > b(p)) \vee (w(p) > d(p))) \quad (2.5)$$

But apply this theorem to the counterexamples. Obviously it is epistemologically possible (perhaps even epistemologically necessary) that disbelieving 'I am in pain', when I am not experiencing pain, is more reasonable than the other two options;

---

<sup>5</sup>Ulm, 1975, 58

<sup>6</sup>Imlay, 1969, 290

(2.5) then entails that it is epistemologically *impossible* for incorrectly believing that I'm experiencing pain to be less reasonable than remaining agnostic. This is an extremely strong claim, and I see no way the defender of Chisholm might argue for it.

### 3. CONSISTENCY I: CEP

If, as I believe, maintaining (2.5) is indefensible, it seems we are left with the possibility that CEP is logically inconsistent. But Chisholm and Keim provide a relative consistency proof<sup>7</sup>: an interpretation of CEP such that all of its axioms are true sentences of the rational numbers. If CEP entailed a logical contradiction, then so would the rational numbers. Since the rational numbers do not entail a logical contradiction, then neither does CEP.

But this is only one sort of consistency, 'inner' or 'formal consistency'. In a short note to himself<sup>8</sup>, Gödel once defined a notion of consistency he called 'outer consistency'. Outer consistency is technically only defined for formal logical systems that prove sentences in the language of elementary arithmetic of a logical form denoted  $\Pi_1^0$ ; a system  $S$  is outer consistent if, and only if, all of these sentences are true theorems of elementary arithmetic. That is, outer consistency is the requirement that a formal system does not prove any false sentences of elementary arithmetic (of a certain logical form). We can expand this notion, at least roughly, to systems like CEP by requiring that they do not prove the negation of any sentences we take to be intuitively true. This makes room for a third interpretation of the counterexamples of the last section: while CEP is inner consistent, the counterexamples show that CEP is not outer consistent, and hence fails as a formal theory of epistemic reasoning.

The defender of Chisholm might say that this is simply begging the question against CEP. Perhaps (2.5) shows that we need to adjust some of our epistemic intuitions. The defender of CEP therefore just needs to give a reasonable or intuitive defence of the axioms of CEP, and (A4) in particular. If this is done, the *prima facie* counter-intuitiveness of (2.5) is turned, and the burden is placed on the critic

<sup>7</sup>Chisholm & Keim, 1972

<sup>8</sup>Gödel, 1972/1986

of CEP instead: why should we accept his analysis of the counterexample situations instead of the analysis given by (A4)?

Chisholm's terms of epistemic appraisal might be able to give us such a intuitive defence of (A4). In particular, as Chisholm and Keim point out<sup>9</sup>, (A4) is equivalent to

$$\forall p (A(p) \rightarrow F(p)), \quad (3.1)$$

if believing  $p$  is epistemically acceptable, then believing  $p$  has some presumption in its favour. (A4) therefore expresses one of the basic relations in Chisholm's hierarchy of the terms of epistemic appraisal, where being acceptable is the midpoint between the high praise of being beyond reasonable doubt and the low praise of merely having some presumption:

“Acceptable,” then, expresses less praise than does “reasonable.” But it expresses more praise than does the doubtful compliment, “believing is more reasonable than disbelieving,” which tells us merely that the proposition has *some* presumption in its favour.<sup>10</sup>

But what reason do we have to accept Chisholm's epistemic hierarchy? In particular, does being acceptable really express higher praise than having some presumption in its favour? Formally, there is little difference between the two terms of praise:

$$\begin{aligned} A(p) &=_{\text{df}} b(p) \geq w(p), \\ F(p) &=_{\text{df}} b(p) > d(p). \end{aligned}$$

There's certainly nothing here that suggests the former ought to imply the later, so perhaps there are intuitive grounds for this hierarchy instead. Thus consider Chisholm's reasons for ranking having some presumption so low:

For in saying that believing is more reasonable than disbelieving we may be saying only that the former is the lesser evil, epistemically. Consider, for example, the proposition that the Pope will be in Rome on the third Tuesday in October five years from now. Believing it,

<sup>9</sup>Chisholm & Keim, 1972, 103

<sup>10</sup>Chisholm, 1977, 10, his emphasis

given the information we now have, is more reasonable than *disbelieving* it, ie, it is more reasonable to believe that the Pope will be in Rome at that time than it is to believe that he will *not* be there. But withholding the proposition, surely, is more reasonable still.<sup>11</sup>

This is a fairly reasonable line of thought: the fact that believing is better than disbelieving tells us nothing about withholding, and indeed withholding might be the best of the three options.

However, it applies, *mutatis mutandis*, to acceptability as well. The fact that believing is better than withholding tells us nothing about disbelieving, and indeed disbelieving might be the best of the three options. Furthermore, we can paraphrase Chisholm's own reasoning against acceptability:

In saying that believing is more reasonable than withholding we may be saying only that the former is the lesser evil, epistemically. Consider, for example, the proposition that  $2 + 2 = 5$ . Believing it, given the information we now have, is at least as reasonable as withholding belief. But disbelieving this proposition, surely, is more reasonable still.

So neither acceptability nor having some presumption should be very high terms of epistemic praise, and there is no reason to think one is any higher than the other.

Attempting to defend (A4) will require a more involved investigation.

#### 4. AVOIDING ERROR AND GETTING TRUTH

Chisholm quotes William James:

There are two ways of looking at our duty in the matter of opinion – ways entirely different, and yet ways about whose difference the theory of knowledge seems hitherto to have shown very little concern. *We must know the truth:* and we *must avoid error* – these are our first and great commandments as would-be knowers; but they are not two

---

<sup>11</sup>Chisholm, 1977, 8, his emphasis

ways of stating an identical commandment, they are two separable laws.<sup>12</sup>

Call these two 'epistemic commandments' Avoid Error and Get Truth. As a crude first approximation, suppose that Avoid Error says that believing  $p$ , when  $p$  is actually false, is always the worst possible option. Whether disbelieving  $p$  or withholding  $p$  is best, believing  $p$  is the worst, and hence withholding  $p$  is more reasonable than believing  $p$ . On the other hand, if  $p$  is actually true, then withholding  $p$  is more reasonable than disbelieving  $p$ . Hence, for any  $p$ , either withholding it is more reasonable than believing it or withholding it is more reasonable than disbelieving it. Similarly, we might suppose that Get Truth says that we should always go ahead and take a chance on belief or disbelief, whatever warrant we have, because we might get lucky. Hence, both believing  $p$  and disbelieving  $p$  are more reasonable than withholding. Symbolically, these become the following<sup>13</sup>:

**Avoid Error:**  $\forall p ((w(p) > b(p)) \vee (w(p) > d(p)))$

**Get Truth:**  $\forall p ((b(p) > w(p)) \wedge (d(p) > w(p)))$

Note first that Get Truth and Avoid Error cannot both be true; indeed, Get Truth asserts that every  $p$  is such that either  $p$  or  $\neg p$  is a counterexample to Avoid Error. This is no great failing, because Get Truth is already so strong that it is *prima facie* implausible. But Avoid Error is also incompatible with the existentially-quantified weakening of Get Truth, ( $\exists T$ ), as this asserts the existence of at least one counterexample. This is one form of the tension at the heart of James' pair of 'epistemic commandments': Avoid Error cannot be true if Get Truth – or some similar, weaker statement – is true. Hence Avoid Error doesn't just say that *sometimes* avoiding error is more important than getting the truth, but that avoiding error is *always* more important than having any truth beliefs at all. It absolutely precludes the second of James' 'epistemic commandments', and is therefore ridiculously strong.

Thus far, this only suggests that Avoid Error is intuitively unacceptable. I prove below that Avoid Error is formally equivalent, in  $CEP^-$ , to (A4). And recall that

<sup>12</sup>James, quoted in Chisholm, 1977, 14, emphasis as given in Chisholm

<sup>13</sup>Unlike the 'epistemic commandments', these two sentences – and their related weakenings – need not be interpreted deontologically. Since Chisholm does have a deontic conception of epistemic logic, however, I interpret these sentences deontologically in what follows.

(A4) is equivalent to (3.1), one of the critical links in Chisholm’s hierarchy of epistemic terms. Hence, if Avoid Error was a reasonable statement of the ‘epistemic commandment’ to avoid error, this might provide the intuitive defence for both (A4) and Chisholm’s hierarchy. However, since the statement of Avoid Error is incompatible with Get Truth and ( $\exists$ T), and therefore ridiculously strong, *a fortiori* so is (A4) and Chisholm’s hierarchy. I take this to be a decisive *reductio ad absurdum* of Chisholm’s system of epistemic logic and epistemic hierarchy.

Note that (A4) is not a trivial hypothesis, easily separated from the core of Chisholm’s epistemology. We have already seen that it is asserted in 1972.<sup>14</sup> Twenty-five years later, in the third edition of *Theory of knowledge*, he ranks a belief being ‘epistemically in the clear’ above being merely ‘probable’ in his epistemic hierarchy<sup>15</sup>. The definitions of these terms correspond exactly to being acceptable and having some presumption. And in a footnote, he asserts that a principle equivalent to (3.1) is ‘needed to complete our hierarchy’<sup>16</sup>. A commitment to (3.1), and *a fortiori* (A4), is crucial to Chisholm’s mature thought. The fact that (A4) is equivalent to the ridiculously strong Avoid Error is devastating.

I now turn to the proof of the claim above.

**Theorem.**  $\text{CEP}^- \vdash (\text{A4}) \leftrightarrow \text{Avoid Error}$

*Proof.* For the sake of human readability, we proceed informally. Both parts of the proof can be easily formalised.

$\Rightarrow$  We have already seen this part of the proof; the claim here is

$$\text{CEP}^- \vdash (\text{A4}) \rightarrow \text{Avoid Error}, \quad (4.1)$$

which is equivalent to (2.5). As in the discussion leading up to (2.5), the proof is a *reductio ad absurdum*.

$\Leftarrow$  Suppose Avoid Error and the negation of (A4), ie,

$$\exists p ((w(p) \leq b(p)) \wedge (b(p) \leq d(p))). \quad (4.2)$$

<sup>14</sup>Ulm’s criticism of (A4) was published in 1969, indicating that the axiom actually dates back even further.

<sup>15</sup>Chisholm, 1989, 16

<sup>16</sup>Ibid, 17n10

Let  $p$  satisfy (4.2). Then, since the first disjunct of Avoid Error is false,  $w(p) > d(p)$ . But, by transitivity and (4.2),  $w(p) \leq d(p)$ , which is equivalent to  $\neg(w(p) > d(p))$ , contradiction.  $\square$

## 5. PRINCIPLE AND CONSTRUCTIVE EPISTEMOLOGY

In 1919, Albert Einstein wrote a short article on the theory of relativity for the London Times, republished the next year in *Science*.<sup>17</sup> In this article, Einstein distinguished between ‘principle’ and ‘constructive’ scientific theories. Constructive theories ‘attempt to build a picture of complex phenomena out of some relatively simple proposition’<sup>18</sup>. Einstein’s example is the kinetic theory of gases, in which the empirically observable properties of gases (the complex phenomena) are explained by postulating that gases are composed of molecules which collide and rebound like billiard balls within a finite-volume container (the relatively simple proposition).

By contrast, theories of principle eschew ‘hypothetical constituents’, and instead start with ‘empirically observed general properties of phenomena, principles from which mathematical formulæ are deduced of such a kind that they apply to every case which presents itself’<sup>19</sup>. Einstein goes on to assert that the special theory of relativity is such a theory of principle. In the special theory from facts about the equivalence of the laws of physics as described within any inertial frame of reference, the isotropy of space and time, and the behaviour of light rays – general principles that are taken to be as empirically well-confirmed as anything – Einstein mathematically derives formal conditions which any physical theory must satisfy, viz, the laws of the theory must have a certain algebraic property called Lorentz covariance. Philosopher of science Yuri Balashov and historian of physics Michel Janssen describe the difference in this way:

In a theory of principle, one starts from some general well-confirmed empirical regularities that are raised to the status of postulates . . . .

With such a theory one explains the phenomena by showing that they necessarily occur in a world in accordance with the postulates.

<sup>17</sup>Einstein, 1920

<sup>18</sup>Ibid, 8

<sup>19</sup>Ibid

Whereas theories of principles are about the *phenomena*, constructive theories aim to get at the underlying *reality*.<sup>20</sup>

Theories of principle are often regarded by metaphysically-minded philosophers of science as non-explanatory. In a Humean fashion, the theory simply raises a set of frequently observed regularities to the status of axioms, and does not make any claims about underlying causes. We can think of them as the data against which a constructive theory is to be tested: if a theory explains electrodynamic phenomena by postulating electromagnetic fields, say, but the properties of these fields are not Lorentz covariant, then we must reject the theory (at least in that form). Alternatively, a principle theory provides a heuristic for the development of a constructive theory. We need to get clear on what general phenomena we are trying to explain – and, say, what algebraic form the explanatory laws must take – before we endeavour to find an explanation.

In the second edition of *Theory of knowledge*, Chisholm identifies ‘two ways of throwing light upon what is intended by an undefined expression. One is to make explicit the basic principles it is used to formulate. The other is to try to paraphrase the undefined expression into a different terminology’<sup>21</sup>. In the first way, we take some expression – say, ‘is more reasonable than’ – as *undefined*, and attempt to illuminate it by showing how to use that expression to define other terms – such as ‘is epistemically acceptable’ and ‘has some presumption in its favour’ – and giving general principles articulating the ways we ordinarily use the expression. It is very much like what logicians and mathematicians call an *implicit definition* of the expression in question. In the second way, we *define* – or *explicitly define* – the expression in terms of other expressions – in the case of reasonability, expressions concerning epistemic duty<sup>22</sup>, or evidence, or reliable faculties, or properly functioning faculties, perhaps.

I’d like to suggest that Chisholm’s two ways track the same division as Einstein’s distinction between principle and constructive theories. Theories of principle correspond to the first, formal way. The axioms used to give some formal theory or

---

<sup>20</sup>Balashov & Janssen, 2003, 331, their emphasis

<sup>21</sup>Op cit, 12-3

<sup>22</sup>Op cit, 14

another in the language  $\mathcal{L}$  are intended to capture the most general and abstract regularities about either the way we use epistemic expressions in ordinary English or our epistemic intuitions (depending on one's philosophical methodology).

This is why both inner and outer consistency are important. Inner consistency shows that our theory does not snag on any contradiction. Outer consistency shows that it accurately captures at least some of our epistemic intuitions. A successful theory of principle must do both things. One of the central claims I have defended in this paper is that, while Chisholm's CEP is inner consistent, it is not outer consistent, and hence fails as a theory of principle.

However, if theories of principle are non-explanatory, and all the various formal theories I have considered here are theories of principle, then none of these theories is in any way epistemically explanatory. None of them can be used to say, for a given agent, why it is that this or that particular belief is actually warranted, or even why this particular belief is more reasonable than that one.

At least, none of them can do this by itself. This sort of work can only be done by a constructive theory, corresponding to Chisholm's second way. Classical phenomenal foundationalism, evidentialism, reliabilism, or proper functionalism, according to the paradigm I'm laying out here, are all constructive theories. Proper functionalism, for example, explains the fact that some belief is more reasonable than another by pointing to facts about the agent's beliefs when she is functioning properly according to her design plan; evidentialism points to facts about the evidence on the basis of which the agent forms the beliefs in question; and so on.

Finally, it is important to recognise that, on the distinction I have been presenting in this section, the direction of fit between principle and constructive theories goes strictly one way. A principle theory is a hermeneutic giving '*a priori*' restrictions on the possibilities for a constructive theory and providing the standard against which a constructive theory is checked. Especially with respect to the latter, to use a constructive theory in developing a theory of principle is at best bootstrapping and at worst downright circular. This is why I have deliberately avoided appealing to any notions tied to particular constructive theories and instead kept my arguments

either formal or intuitive, and likewise why I have avoided building any notions tied to particular constructive theories into my preferred formal system<sup>23</sup>.

Furthermore, as a hermeneutic, constructing a theory of principle is necessary for the successful development of a constructive theory. Without these restrictions, there are too many options, and no solid principles for deciding between two options.<sup>24</sup> Our current best theory of space and time, the constructive General Theory of Relativity, could never have been developed without the Special Theory, a theory of principle. In a Cartesian metaphor, a search for knowledge without a hermeneutic guiding it is aimless, fruitless wandering (only a mildly mixed metaphor) through the forests of knowledge. In a mathematical metaphor, the theory of principle restricts the space of constructive theories to a manageable size.

Contemporary Analytic epistemology seems, from my outsider's perspective, to consist largely of interminable skirmishes over the first principles grounding a plethora of constructive theories. In addition, philosophers of science infamously complain that the bulk of these theories are largely irrelevant to the task of giving an epistemology for natural science. I close this paper by suggesting that the lack of a formal epistemic logic could very well be one major cause of this disordered state of affairs. Contemporary Analytic epistemologists have set off in search of constructive theories without the guidance of a theory of principle, and find themselves lost in the Cartesian woods.

#### APPENDIX A. CONSISTENCY II: OTHER THEORIES

There are at least two plausible rivals to CEP, at least in terms of formal theories in the language  $\mathcal{L}$ . One is the theory where (A4), equivalent to Avoid Error, is replaced with an axiom equivalent to Get Truth. Since, for any  $p$ ,  $d(p) > w(p)$  is

---

<sup>23</sup>Conditional-CEP<sup>+</sup>, described in the appendix.

<sup>24</sup>Yet another way of understanding the theory of principle is that it *just is* the collection of the most general and abstract principles for deciding between two rival constructive theories. However, this should not be understood as the claim that the theory of principle aspires to be a complete catalogue of all intuitions, or to clearly favour one kind of constructive theory over all rival approaches. It is a formal theory, giving only the most general and abstract treatment of the subject matter, and other, non-formal and more particular, principles will have to be brought into play when it comes time to evaluate constructive theories.

equivalent in  $\text{CEP}^-$  to  $b(\neg p) > w(\neg p)$ , we can eliminate the second conjunct of Get Truth, giving us the equivalent sentence

$$(\forall T) \forall p (b(p) > w(p)).$$

As given above in §1.3, the addition of this axiom to  $\text{CEP}^-$  is the formal theory GT.

The third formal theory is one where we moderate both Avoid Error and Get Truth, replacing the universal quantifiers with existential, and add both axioms to  $\text{CEP}^-$ . This is the formal system  $\text{CEP}^+$  given in §1.4. Note that one might want slightly more useful, conditional axioms in place of  $(\exists T)$  and  $(\exists E)$ :

$$(\exists T') \forall p (\tau(p) \rightarrow ((b(p) > w(p)) \wedge (d(p) > w(p))))$$

$$(\exists E') \forall p (\eta(p) \rightarrow ((w(p) \leq b(p)) \rightarrow (b(p) > d(p))))$$

Loosely,  $\tau$  and  $\eta$  give conditions under which Get Truth or Avoid Error are to be applied.<sup>25</sup> Call the version of  $\text{CEP}^+$  that uses these axioms conditional- $\text{CEP}^+$ .

For now, we consider only the following possible ‘defence’ of CEP. Chisholm and Keim, so says the defender, have already shown that CEP is inner or formally consistent. In particular, and as mentioned before, they assign to every doxastic attitude  $b(p)$ ,  $d(p)$ , or  $w(p)$  a rational number  $V'(b(p))$ ,  $V'(d(p))$ , or  $V'(w(p))$ , respectively, subject to the conditions that  $V'(b(p)) = -V'(d(p)) = -V'(b(\neg p))$  and  $V'(w(p)) = V'(w(\neg p)) = \frac{1}{|V'(b(p))|}$  for all  $p$ .<sup>26</sup> In addition, for doxastic attitudes  $\varphi, \psi$ ,  $V'(\varphi) > V'(\psi)$  if, and only if,  $\varphi > \psi$ . The axioms (A1)-(A7) thereby come out as true sentences in the rational numbers. The defender of CEP continues by pointing out that, on this interpretation, Avoid Error is true, while Get Truth,  $(\forall T)$ , and  $(\exists T)$  are false, so this model cannot be used to show that GT and  $\text{CEP}^+$  are formally consistent. This places a certain burden on the advocate for either GT or  $\text{CEP}^+$ .<sup>27</sup>

To discharge this burden, we must develop a model like  $V'$ , which we call  $V$ . Like,  $V'$ ,  $V$  takes doxastic attitudes  $b(p)$ ,  $d(p)$ ,  $w(p)$  to rational numbers  $V(b(p))$ ,  $V(d(p))$ , and  $V(w(p))$ , respectively, subject to the following conditions for all sentences  $p$  of

<sup>25</sup>Strictly speaking, the introduction of  $\tau$  and  $\eta$  requires moving to a new formal language.

<sup>26</sup>Chisholm and Keim do not name the function  $V'$  defining the model.

<sup>27</sup>Conditional- $\text{CEP}^+$  is formally consistent if, and only if,  $\text{CEP}^+$  is.

$\mathcal{L}'$  and doxastic attitudes  $\varphi, \psi$ :

$$\begin{aligned} V(b(p)) &> 0 & V(b(\neg p)) = V(d(p)) &= \frac{1}{V(b(p))} \\ V(w(p)) = V(w(\neg p)) &= -V(b(p)) - \frac{1}{V(b(p))} & V(\varphi) > V(\psi) &\leftrightarrow \varphi > \psi \end{aligned}$$

These conditions may strike the reader as rather silly. We certainly cannot use  $V$  as a way of somehow evaluating the reasonableness of our doxastic attitudes, unless withholding is always the worst possible option, that is, unless Get Truth is true. But this objection gets the purpose of  $V$  wrong.  $V$  is intended *only* to show that GT and CEP<sup>+</sup> are *inner* consistent. Putting conditions on the class of models for the language  $\mathcal{L}$  – requiring models to not conflict with our epistemic intuitions – is imposing conditions on the *outer* consistency of any formal theory constructed in  $\mathcal{L}$ . Like Hilbert’s system of love, law, chimney-sweep, we can freely make use of counter-intuitive models when checking formal or inner consistency.

As I have already argued, outer consistent conditions are entirely reasonable, and indeed should be decisive in adopting one formal theory or another as ‘the’ formal theory of the subject matter, but they go beyond both the scope of Chisholm’s presentation of CEP and this appendix. Indeed, my purpose here is to show that the defender of CEP cannot dismiss the rival theories on grounds other than outer consistency.

Next we check the axioms (A1)-(A7) (except of course (A4)), ( $\forall T$ ), ( $\exists T$ ), and ( $\exists E$ ). Axioms (A1) and (A2) are ordinary order axioms satisfied by the greater-than relation on the rationals. For the other axioms, we use variables  $x, y$  for positive rationals. We will copy each axiom, transcribe it as an open-variable sentence in the language of the rational numbers, and include notes on why the resulting sentence is true.

$$\begin{aligned} \text{(A3)} \quad \forall p, q \quad ((b(p) > b(q)) &\leftrightarrow (d(q) > d(p))) \\ x > y &\leftrightarrow \frac{1}{y} > \frac{1}{x} \end{aligned}$$

This holds since both  $x$  and  $y$  are positive.

( $\forall T$ )  $\forall p (b(p) > w(p))$

$$x > -x - \frac{1}{x}$$

Since  $x > 0$ ,  $0 > -x - \frac{1}{x}$ , and by transitivity the transcribed sentence holds.

( $\exists T$ )  $\exists p ((b(p) > w(p)) \wedge (d(p) > w(p)))$

This sentence is entailed by ( $\forall T$ ) if there is at least one  $p$  such that  $V(b(p)) > 0$ , and hence will be true in any non-trivial model.

( $\exists E$ )  $\exists p ((w(p) \leq b(p)) \rightarrow (b(p) > d(p)))$

For some  $x$ , if  $-x - \frac{1}{x} \leq x$ , then  $x > \frac{1}{x}$ .

$x$  satisfies the sentence if, and only if,  $x > 1$ , and hence will be true in any non-trivial model.

(A5)  $\forall p, q ((w(p) = w(q)) \leftrightarrow ((b(p) > b(q)) \vee (d(p) > b(q))))$

$$\left( \frac{x^2 + 1}{x} = \frac{y^2 + 1}{y} \right) \leftrightarrow \left( x = y \vee \frac{1}{x} = y \right)$$

$\boxed{\leftarrow}$  If  $x = y$ , then the left hand side is obviously true. If  $xy = 1$ , then the left hand side becomes

$$x^2y + y = y^2x + x \tag{A.1}$$

$$x + y = y + x, \tag{A.2}$$

which is obviously true.

$\boxed{\rightarrow}$  Again, the left hand side is

$$x^2y + y = y^2x + x. \tag{A.3}$$

We treat this as a quadratic equation of the variable  $x$ , and apply the quadratic formula.

$$yx^2 - (y^2 + 1)x + y = 0 \quad (\text{A.4})$$

$$x = \frac{y^2 + 1 \pm \sqrt{(y^2 + 1)^2 - 4y^2}}{2y} \quad (\text{A.5})$$

$$= \frac{y^2 + 1 \pm \sqrt{(y^2 - 1)^2}}{2y} \quad (\text{A.6})$$

$$= \frac{y^2 + 1 \pm (y^2 - 1)}{2y} \quad (\text{A.7})$$

$$= y \text{ or } \frac{1}{y}, \quad (\text{A.8})$$

as desired.

$$(\text{A6}) \quad \forall p, q \quad (((b(p) > b(q)) \wedge (b(p) > d(q))) \rightarrow (w(q) > w(p)))$$

$$\left( x > y \wedge x > \frac{1}{y} \right) \rightarrow \left( \frac{y^2 + 1}{y} < \frac{x^2 + 1}{x} \right)$$

Set  $z = x - y$ . By the first conjunct of the antecedent,  $z > 0$ . Then we want to show that, if  $xy > 1$ ,

$$yx^2 - (y^2 + 1)x + y > 0. \quad (\text{A.9})$$

That is,

$$y(y + z)^2 - (y^2 + 1)(y + z) + y = y(y^2 + 2zy + z^2) - (y^3 + zy^2 + z + y) + y \quad (\text{A.10})$$

$$= y^3 + 2zy^2 + z^2y - y^3 - zy^2 - z - y + y \quad (\text{A.11})$$

$$= zy^2 + z^2y - z > 0. \quad (\text{A.12})$$

Now, this holds if, and only if,

$$zy^2 + z^2y > z. \quad (\text{A.13})$$

And since  $z > 0$ , this is equivalent to

$$y^2 + zy > 1. \tag{A.14}$$

Substituting back in the definition of  $z$ , the left hand side here is

$$y^2 + (x - y)y = y^2 + xy - y^2 \tag{A.15}$$

$$= xy. \tag{A.16}$$

That is, given that  $x > y$ , the conclusion holds if, and only if,  $xy > 1$ , which is the second conjunct of the antecedent. Hence, the sentence holds in the model.

$$(A7) \quad \forall p (w(p) = w(\neg p))$$

$$x + \frac{1}{x} = \frac{1}{x} + x$$

Thus, both GT and CEP<sup>+</sup> are at least formally or inner consistent. The choice between CEP, GT, and CEP<sup>+</sup> must be based strictly on which theory best expresses our epistemic intuitions – that is, which, if any, is outer consistent. And here CEP<sup>+</sup> clearly comes out on top, as it is the only one of the three capable of accommodating both of James' ‘epistemic commandments’. Conditional-CEP<sup>+</sup>, where axioms (∃T) and (∃E) are replaced with axioms that include conditions  $\tau$  and  $\eta$ , is still more promising.

## REFERENCES

- Balashov, Y., Janssen, M. (2003). Presentism and relativity. *British journal of the philosophy of science*, 54, 327-46.
- Chisholm, R. (1977). *Theory of knowledge* (second ed.). Prentice-Hall.
- Chisholm, R. (1982). *The foundations of knowing*. University of Minnesota.
- Chisholm, R. (1989). *Theory of knowledge* (third ed.). Prentice-Hall.
- Chisholm, R., Keim, R. (1972). A system of epistemic logic. *Ratio*, 14(2), 99-115.
- Einstein, A. (1920). Time, space, and gravitation. *Science*, 51, 8-10.
- Gödel, K. (1986). Some remarks on the undecidability results. In S. Feferman, J. Dawson, S. C. Kleene, G. H. Moore, R. M. Solovay, J. van Heijenoort

- (Eds.), *Collected works, vol. II* (p. 281-306). Oxford. (Original work published 1972)
- Imlay, R. (1969, December). Chisholm's epistemic logic. *Philosophy and phenomenological research*, 30(2), 290-3.
- Ulm, M. (1975, January). Chisholm's fourth axiom. *Philosophical studies*, 27(1), 57-61.