

## Nozick's fourth condition

### *Introduction*

Nozick's tracking account of knowledge includes four individually necessary and jointly sufficient conditions. S knows  $p$  iff (i)  $p$  is true; (ii) S believes  $p$ ; (iii) if  $p$  were false S would not believe  $p$ ; (iv) if  $p$  were true, then S would believe  $p$ .<sup>1</sup>

In this paper I shall argue that the fourth condition, which we will call Adherence<sup>2</sup>, cannot be a necessary condition on knowledge. The argument proves that the following:

(AD) S knows  $p \square (p \square \rightarrow S \text{ believes } p)$ <sup>3</sup>

is inconsistent with the possibility of a *perfectly cautious believer*. A perfectly cautious believer is one who is sufficiently careful and reluctant to take risks with believing, that he is reliably disposed to believe only when by believing he comes to know. The perfectly cautious believer is someone of whom the following is true:

(PCB) S believes  $p \square S$  knows  $p$

The perfectly cautious believer may be ideal (and may have other faults we shall come to). But he is not incoherent. It cannot be a mere condition on the concept of knowledge that a perfectly cautious believer is impossible. Consequently (AD) is false and Adherence, Nozick's fourth condition, cannot be part of the analysis of knowledge.

### *Adherence*

Nozick intends that the counterfactual in (AD) needs to be understood in a way that is stronger than the reading yielded by Lewis's account. According to Lewis,  $p \square \rightarrow q$ , is true if at the nearest possible world in which  $p$  is true, and at those equally close,  $q$  is also true. Consequently in a world where  $p$  is true and  $q$  is true,  $p \square \rightarrow q$  is also true. But according to Nozick,  $p \square q$  does not entail

---

<sup>1</sup> Robert Nozick 1981 *Philosophical Explanations* Oxford: Clarendon Press; 172-8.

<sup>2</sup> C.f. A.R.Lacey 2001 *Robert Nozick* Chesham: Acumen; 103.

<sup>3</sup> Nozick does amend Adherence, in order to exclude cases of subjects who have contradictory beliefs, to:  $p \square \rightarrow S$  believes  $p$  and not-(S believes not- $p$ ); Nozick 178. This amendment is irrelevant to what follows.

$p \Box \rightarrow q$ . When  $p$  is true it is required not only that  $q$  is true but also that  $q$  is also true at all close worlds where  $p$  is true.

Now let us turn to the application of this conditional to Nozick's account to knowledge. Since knowledge entails truth, as required in Nozick's first, uncontentious, condition, the cases we are interested in are precisely those where the antecedent of the conditional is true. Hence Adherence may be expressed as:

(AD\*) S knows  $p \Box$  S believes  $p$  in all close worlds where  $p$  is true.

Nozick includes Adherence in order to account for two kinds of case not captured by the remaining three conditions. One kind of case is exemplified by the well-known story from Harman.<sup>4</sup> The death of the dictator is initially reported in the media but the reports are shortly thereafter and convincing denials are widely issued by the government. Smith is one of the few who hears and believes the initial reports but by chance fails to hear the later denials. Smith satisfies the other conditions ((i)-(iii) above), since he has a true belief that he would not have had were it false (were the dictator still living, there would have been no initial report). Because he could so easily not have had the true belief, we are disinclined to count this as knowledge. Adherence does the job of ruling this case out as knowledge, since there are nearby worlds (where Smith hears the later denials) in which Smith does not believe that the dictator is dead, even when he is dead. The other kinds of case concerns failure to know necessary truths. One can have a belief in a necessary truth that fails to be knowledge, but such a belief will satisfy (iii) as well as (i) and (ii), since the antecedent of the counterfactual in (iii) (were  $p$  false) cannot be satisfied, the counterfactual is trivially true. Someone who believes a necessary truth for a bad reason (gullible faith in an unreliable source) will nonetheless fail to satisfy Adherence, since there will be nearby worlds in which though  $p$  is true, the subject does not believe  $p$  (had the unreliable source said something different, the subject would not have believed  $p$ ).<sup>5</sup>

### *The disposition to believe and the capacity to know*

Roughly speaking, Adherence suggests that the knower has a certain disposition to believe. For the knowing subject not only believes in the actual world but in certain close worlds. Strictly speaking

---

<sup>4</sup> Gilbert Harman 1973 *Thought* Princeton: Princeton University Press; 142-154. C.f. Nozick 1981; 177.

<sup>5</sup> Nozick 1981; 186-7.

this need not in every case be the result of a disposition of the knower's—after all the counterfactual could be made true by a (benevolent?) demon who is prepared to plant true beliefs in the subject's head. The latter fact itself might be thought to be an objection to Adherence, although the matter is not especially clear. There are other problems, since one can come to know by chance—I see a television new item while randomly flicking through the channels. In a nearby world I do not chose that channel at that moment and so fail to believe the relevant proposition.<sup>6</sup> Such problems are addressed (if not necessarily entirely resolved) by employing the idea of a method by which one comes to believe, which must remain constant over the relevant possible worlds.<sup>7</sup>

In what follows I shall argue that Adherence cannot be right since it requires the disposition to believe to exceed the capacity to know. Even if odd instances that might instantiate Adherence do not involves dispositions to believe, standard cases do. For example, if a normally sighted subject looks at a bright blue swatch of material in good conditions and thereby comes to know that the swatch is blue, it is true that the subject would believe the swatch is blue in slightly different conditions in which the swatch is still blue (e.g. the swatch is slightly less bright or more bright). That counterfactual is made true by a disposition (or set of related dispositions) of the subject, such as the disposition to believe of things of that are that shade or similar, when presented in good conditions, that they are blue.

(AD\*) says that if S knows  $p$  in  $w$  then S believes  $p$  in close  $p$ -worlds (worlds where  $p$  is true). Does S know  $p$  in all these worlds? If not, then S must be disposed to believe in worlds where S will not know  $p$ ; that is S's disposition to believe outstrips his capacity to know. If, on the other hand, S does know  $p$  in all the close  $p$ -worlds, then by (AD\*), S believes  $p$  in all  $p$ -worlds close to these  $p$ -worlds. Once again, if belief does not outstrip knowledge, then S knows  $p$  in all these worlds too. And so on. The conclusion is that either S has a disposition to believe that outstrips knowledge or S knows  $p$  in all  $p$ -worlds close to the actual world, and all  $p$ -worlds close to those worlds, and so on. Let us look at this consequence of Adherence in more detail. Let two worlds be  $\text{close}_p$  iff they are close and  $p$  is true in both. And let  $\text{close}_p^*$  be the ancestral of the  $\text{close}_p$  relation. (I.e.  $w_0$  and  $w_n$  are  $\text{close}_p^*$  iff there is some sequence of possible  $p$ -worlds,  $w_1 \dots w_j$  such that (i) every member of the sequence is close to its successor, (ii)  $w_0$  is close to  $w_1$ , and (iii)  $w_j$  is close to

---

<sup>6</sup> This is analogous to the case of the bystander who knows that the bank robber is Jesse James because the latter's mask slipped accidentally and briefly. Nozick 1981; 193.

<sup>7</sup> Nozick 1981; 179-185.

$w_n$ .) Let  $C_{p,w}$  be the set of worlds close $_p^*$  to  $w$ . The conclusion of this argument can be expressed thus:

Adherence (AD\*)

*entails*

either a subject S's disposition to believe outstrips his capacity to know  
or if S knows  $p$  in a world  $w$  then S also knows  $p$  in every world in  $C_{p,w}$ .

But the second disjunct in the entailed proposition above is highly unlikely to be fulfilled.  $C_{p,w}$  can be an extremely large set of worlds. Since the close relation is not transitive neither is close $_p$ . Hence two worlds might well be close $_p^*$  without being at all close (or close $_p$ ). A sequence of worlds ( $w_0$  to  $w_n$ ) may be such that every one differs from its predecessor only by a small variation in circumstances (the shade of a swatch is slightly less blue and slightly more green) yet the first and last world differ markedly (in the first the swatch is blue (B) in the last it is green (G)). One may know some proposition in the first world,  $w_0$ , that one cannot know in the last world,  $w_n$ , even if that proposition is true in all the worlds. Here is an example. For example  $p$  is some proposition true in all the worlds in the sequence; hence the following are also true: i.e. if the swatch is blue then  $p$ , ( $B \square p$ ), if the swatch is green then  $p$ , ( $G \square p$ ). Let us be that in all the worlds S has the same background knowledge, that if the swatch is blue then  $p$  ( $B \square p$ ) but in none of the worlds does S know ( $G \square p$ ). Hence in  $w_0$  S can come to know  $p$  (by modus ponens on the known propositions B and  $B \square p$ ), but S cannot know  $p$  in  $w_n$  (because S does not know  $G \square p$ ). Hence knowledge of  $p$  will not always be ubiquitous in  $C_{p,w}$  even when S knows  $p$  in  $w$ .<sup>8</sup>

And so if (AD\*) is true the disposition to believe exceeds the capacity to know. By this is meant that the subject must have a disposition to believe  $p$  in circumstances in which the subject would not know  $p$ . Returning to our example, let  $w_k$  be a world where S knows B and infers and comes to know  $p$ . Since S knows  $p$  in  $w_k$  (AD\*) requires S must believe  $p$  in  $w_{k+1}$ . Since S's knowledge and beliefs concerning  $p$  come only by inference from ( $B \square p$ ) and B, we may take it that if S believes  $p$  in  $w_{k+1}$ , then that is because S believes B in  $w_{k+1}$ .

---

<sup>8</sup> It may be noted that this case involves inference and so may be thought to fall foul of Nozick's rejection of closure. Two comments on this. First, the failures of closure occur in cases where the inferred proposition is the negation of a remote and sceptical possibility (which means that condition (3) might fail for such propositions). Secondly, the proof given below of the failure of Adherence does not involve a subject making inferences between beliefs and so closure does not come into question.

Imagine that S in believing B and believing  $p$  in  $w_{k+1}$  thereby *knows* B and knows  $p$  in  $w_{k+1}$ . Then (AD\*) tells us that S believes B and believes  $p$  in  $w_{k+2}$ . However, since S cannot know  $p$  in all the worlds in question, there must be a last world in the sequence in which S knows  $p$ . Call this  $w_f$ . Then by (AD\*) S believes  $p$  in  $w_{f+1}$ . Since  $w_f$  is the last world in the sequence in which S knows  $p$ , S does not know  $p$  in  $w_{f+1}$ . Hence in  $w_{f+1}$  S believes a proposition S does not know. The world  $w_{f+1}$  differs from  $w_0$  in the shade of the swatch and in consequent occurrent mental states S has (S's visual impression of the swatch and what S knows). The worlds do not differ in S's dispositions to believe. Hence the disposition to believe manifested in  $w_{f+1}$  is a disposition S possesses in  $w_0$ . Similarly the capacity to know, manifested in  $w_f$  but not in  $w_{f+1}$ , is a capacity S possesses in  $w_0$ . (Furthermore, we may suppose that in all the worlds that S exercises his judgment concerning whether B to the best of his abilities. So that the failure to know in  $w_{f+1}$  is a failure of S's capacity to know, not a piece of bad luck.) What, in  $w_0$ , S is disposed to believe exceeds what S is capable of knowing.

The argument I present below takes issue with this consequence of Adherence. Adherence, remember, is supposed to state a condition that is part of the analysis of the concept of knowledge. But why should it be part of the concept of knowledge, that one's disposition to believe should exceed one's capacity to know? Is it not possible that a knowing subject should be a cautious believer, who is reluctant to believe except in circumstances where knowledge is certain and so is not disposed to believe in circumstances where he would not know? Does not a rational subject want to get his dispositions to believe to match his capacities to know and not to exceed them?

### *Perfectly cautious believers*

I claim that the notion of a perfectly cautious believer is coherent. A perfectly cautious believer is one whose disposition to believe does not exceed his capacity to know, who therefore believes only when by believing he can come to know. Such an individual might be one who doubts his cognitive capacities and so exercises them only in circumstances that ensure their success. A limiting case of a perfectly cautious believer would be a sceptic of the stoic school who suspends belief on all matters. Such an individual is of little interest to us, since the argument concerns those who do have some beliefs and so some knowledge. The latter kind of subject is one who, for example, exercises perceptual judgment only when lighting conditions are good. It might nonetheless be the case that this individual's eyesight is sufficiently satisfactory that his perceptual judgments would

be knowledge even when exercised in cases of moderate lighting conditions. Consider a set of world  $w_0$  to  $w_n$ , for large  $n$ , satisfying the following conditions:

- (i) in  $w_0$  S is looking in the direction of a red apple, which is three feet away, and lighting conditions are excellent;
- (ii) each  $w_{i+1}$  is just like  $w_i$ , except that lighting conditions are slightly poorer;
- (iii)  $w_n$  is like  $w_0$ , except that lighting conditions are extremely poor (it is pitch dark).

Consider the following proposition,  $p$ : ‘there is a red apple near me’. In  $w_0$  S believes and knows  $p$ , since lighting conditions are excellent. But in  $w_n$  S neither believes nor knows  $p$ , since he can see nothing at all. The following is a proof that Adherence and the assumption that S is a perfectly cautious believer are inconsistent. the proof proceeds by a simple mathematical induction:

- (AD\*) In all worlds, S knows  $p \iff$  S believes  $p$  in all close worlds where  $p$  is true.
- (PCB) In all worlds, S believes  $p \iff$  S knows  $p$
- (assumption 1) In  $w_0$  S knows  $p$
- (assumption 2) In  $w_n$  S does not know  $p$
- (assumption 3) In all  $w_i$  ( $0 \leq i < n$ )  $p$  is true
- (assumption 4) For all  $w_i$  ( $0 \leq i < n$ )  $w_i$  is close to  $w_{i+1}$
  
- (base step) In  $w_0$  S knows  $p$  (from assumption 1)
  
- (induction step) Let S know  $p$  in  $w_i$ . By (assumption 3)  $p$  is true in  $w_{i+1}$  and by (assumption 4)  $w_{i+1}$  is close to  $w_i$ . By (AD\*), S believes  $p$  in  $w_{i+1}$ . By (PCB) S knows  $p$  in  $w_{i+1}$ . Hence for all  $w_i$  ( $0 \leq i < n$ ) if S knows  $p$  in  $w_i$  then S knows  $p$  in  $w_{i+1}$ .
  
- (conclusion) In all  $w_i$  ( $0 \leq i < n$ ) S knows  $p$ . Hence in  $w_n$  S knows  $p$ .

The conclusion contradicts assumption 2.

The assumptions 1-4 are merely descriptions of the sequences of possible worlds under consideration and are not contentious. The contradiction is generated by the combination of (AD\*)

and (PCB). Since (PCB) is coherent, the fault must lie with (AD\*). (AD\*) is simply an expression of Nozick's Adherence. Hence Adherence is false.

Defenders of Adherence may wish therefore to put pressure on my contention that (PCB) is coherent. It is worth considering the dialectic here. Perhaps we will find something wrong with (PCB). But if Adherence really were part of the concept of knowledge the above proof shows that that we can know that (PCB) is false just by *a priori* reasoning from the concept of knowledge. That seems to be implausible. But in any case that does require that for a defence of Adherence the objections to (PCB) must relate somehow to the concept of knowledge. For example, it cannot be a relevant objection that (PCB) is an implausible assumption concerning human nature or that it is at best an unrealisable idealisation. Arguably, it is relevant that a subject of whom (PCB) is true is one who is not susceptible to deception by an evil demon or by envatment. But note that for the above proof I do not need S to be a perfectly cautious believer across the board (i.e. in all possible worlds, with respect to all beliefs S might have). S might be *locally* perfectly cautious. In particular all I need is that S be cautious in the range of possible worlds considered in the argument, with respect to the proposition *p*. (I.e. in the proof, (PCB) can be replaced by 'in all  $w_i$  ( $0 \leq i < n$ ) S believes *p*  $\square$  S knows *p*'.) Even if S could be deceived by a demon, it could also be true that in demonless, normal worlds, S is perfectly cautious with respect to perceptual beliefs of the kind being considered.

A different objection suggests that someone of whom (PCB) is true is subject to a certain kind of irrationality. In our example S is someone whose eyesight is good enough to permit him to know *p* even in moderate lighting conditions, although S's caution leads S not to believe in those conditions. So S desists from believing in conditions where S could come to know. Is not that some kind of failure of rationality? I am inclined to agree that it is (other things being equal). This objection has no force against the (ideal) individual who, by whatever means, has managed to match perfectly her disposition to believe to her capacity to know. But even as regards those who do suffer from the irrationality of caution in circumstances where knowledge is possible, that irrationality that does not impact on the coherence of (PCB). (PCB) does not assert that the subject is rational. Irrationality is possible (and actual). And so this objection cannot require us to deny (PCB). Perhaps, however, the charge of irrationality means that (PCB) is inconsistent with some other assumptions made about the subject. In particular, perhaps the irrationality in question is inconsistent with the assumption that S knows *p* in  $w_0$ —on the grounds that in general knowledge and irrationality are inconsistent. But there is no reason to suppose that to be the case. A subject

may show some degree of irrationality in some areas yet come to know things; if that were not so, we would none of us know anything. The mild irrationality of failing to believe in some circumstances that would permit one to know does not eliminate all possibility of knowledge. (Indeed, supposing that it does would undermine the charge that the individual is irrational in failing to believe what he *could* know.)

### *Alternatives to Adherence*

If Adherence is false, what should replace it? If that question means, what should Nozick have as a fourth condition to account for the cases not captured by conditions (i)-(iii) (e.g. Harman's case), then the question assumes that we want to work within a Nozickian framework. But the core of the account, condition (iii) has come under sufficient criticism that we might want to abandon this approach altogether. It might turn out that another general approach will cover these particular cases without an explicit condition included for that specific purpose. For example, a reliabilist account might exclude Harman's case on the grounds that the circumstances are such that Smith's belief-forming process (believing media and government reports in this dictatorship) is unreliable, (even when it generates a true belief). Similarly, reliabilism can deal straightforwardly with the case of necessary truths believed for bad reasons. Or we might agree with Williamson that the failure of Nozick's analysis, along with the failure of every other analysis of knowledge that has been proposed, is evidence that there is no analysis to be had at all, and so no urgent need to find a replacement for Adherence.

At the same time, there remains a question mark over Harman's case does fail to be knowledge, and hence over whether we need anything in place of Adherence. It is not immediately obvious that Harman's Smith lacks knowledge. That case is a case of someone who has a true belief for good reasons. However there is nearby misleading counter-evidence that the subject by chance fails to notice, and which would have caused the subject to drop his belief, had he noticed it. And indeed Nozick himself registers some doubts in connection with such cases. For example, he mentions the following case. You truly believe (as a result of what you see) that Tom has removed a book from the library. However a few yards away, but by chance unheard by you, Tom's mother is telling someone that Tom is out of town and that Tom has an identical brother. Unbeknownst to all the mother is a pathological liar. Nozick himself says that it is unclear whether you know or do not know that Tom took the book. He also says that in such cases it is unclear whether Adherence holds or not, which would mean that the unclarity in the concept of knowledge is matched by a

corresponding unclarity in the analysis. However, if this is not a clear case of not knowing, why should Harman's case be a clear case of not knowing? And if the latter is not such a clear case, then our motivation for wanting Adherence or a replacement for it is not very strong. Indeed, it seems that there are cases with the structure in question (nearby misleading but unnoticed evidence) which do yield knowledge (and so constitute counterexamples on their own to Adherence). Ishbel is a scientist who is investigating the claim that there is a causal connection between Fs and Gs in some population (e.g. smoking and cancer). Having sampled the population as much as she can, she compares the proportion of Fs that are Gs with the proportion of non-Fs that are Gs. The former proportion is greater and Ishbel applies the appropriate statistical techniques. These techniques tell her that the null hypothesis (that there is no relationship between Fs and Gs) is falsified. However, had the proportion of Fs that were also Gs been very slightly smaller, those techniques would not have told her that the null-hypothesis is refuted (she would have needed more data to come to a judgment). Now since there are many Fs that are not Gs around (e.g. cancer-free smokers) Ishbel might easily have sampled some Fs that are not Gs where in fact she sampled Fs that are Gs. That is, she might easily have gathered evidence that would have prevented her from drawing the conclusion (and hence gaining the knowledge) that the null-hypothesis is false. But we would not want to rule this out as a way of getting knowledge. Furthermore, we might apply very rigorous statistical tests in order to avoid false positives. Because a test is rigorous many of the cases that pass the test pass it only just. Because it is rigorous it does yield knowledge when a case does pass it, however narrowly. But in the cases of a narrow pass, it could have occurred that with very slightly different evidence the test would have been failed, and the investigator would have withheld judgment. Such cases seem to be cases of knowledge where one could easily not have had the belief in question (even if true nonetheless).

This brings us back to the central argument of this paper. Adherence requires that for knowledge belief must be a fairly robust tracker of the truth—belief should remain in nearby cases where the belief is true. Consequently, according to Adherence, knowledge is incompatible with nearby contrary (even if misleading) evidence *when knowledge of that evidence would make one give up one's belief*. Note that adherence doesn't make knowledge incompatible with nearby contrary evidence on the grounds that one's belief would be unjustified if one had that evidence. The latter might be what rules out knowledge in some cases. But it is a different requirement from Adherence. Adherence requires that the subject not be disposed easily not to believe. Adherence can fail to be satisfied when there is enough evidence around that would make a belief unwarranted (and so a rational believer might easily not have believed). But it can also easily fail to be satisfied

by a subject who is especially sensitive even to small amounts of contrary evidence. In the latter case Adherence may fail even though the contrary evidence is not enough to make a belief unjustified. The cautious believer is an instance of this—for example a subject who applies very high standards in drawing inferences from the evidence. A cautious believer is disposed to desist from judgment, even though the judgment would have been justified. Often, therefore, the cautious believer might not have believed. In such cases Adherence entails that there is no knowledge.

This is a perverse conclusion. Compare A and B:

A has very high standards for judgment. Proposition  $p$  meets these standards and A believes  $p$ . But  $p$  only just meets A's high standards, and with slightly different (and easily possible) evidence,  $p$  would not have met the high standards and A would not have believed  $p$ .

B has laxer standards for judgment. On the basis of the same evidence as in A's case,  $p$  meets B's standards easily, and B believes  $p$ . Since B's standards are lax, even with the slightly different evidence, B would still have believed  $p$ .

Assuming that B's standards are satisfactory, Adherence means that A does not know  $p$  but B does. This is clearly wrong. Someone who judges on the basis of high standards cannot fail to know what someone who makes the same judgment with the same evidence using laxer standards succeeds in knowing.

### *Conclusion*

Nozick's fourth condition, Adherence, tells us that to have knowledge a subject must be disposed to believe in nearby worlds where the proposition known is true. By considering whether or not the subject not only believes in these worlds but also knows, we reached the conclusion that at some world the subject must know  $p$  while in nearby worlds he believes  $p$  but without knowing. That is, Adherence requires that a subject's disposition to believe must outstrip his capacity to know. But that cannot be a condition on knowledge. It is perfectly possible to be an individual who is cautious about believing, whose capacity to know extends as far as his or her disposition to believe (and perhaps exceeds it in the sense that were she to believe she would know). It is no part of the concept of knowledge that a knowing cautious believer cannot exist.

So Adherence is no part of the concept of knowledge. I have suggested that we do not need to replace it with an improved condition. Other, non-Nozickian approaches can handle the relevant cases just as well, and indeed there is doubt whether some of them need handling at all.