KNOWLEDGE A Human Interest Story

BRIAN WEATHERSON



https://www.openbookpublishers.com

©2024 Brian Weatherson



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). This license allows you to share, copy, distribute and transmit the text; to adapt the text for non-commercial purposes of the text providing attribution is made to the author (but not in any way that suggests that they endorse you or your use of the work). Attribution should include the following information:

Brian Weatherson, *Knowledge: A Human Interest Story*. Cambridge, UK: Open Book Publishers, 2024, https://doi.org/10.11647/OBP.0425

Further details about the CC BY-NC license are available at http://creativecommons.org/licenses/by-nc/4.0/

All external links were active at the time of publication unless otherwise stated and have been archived via the Internet Archive Wayback Machine at https://archive.org/web

Any digital material and resources associated with this volume will be available at https://doi.org/10.11647/OBP.0425#resources

ISBN Paperback: 978-1-80511-394-2 ISBN Hardback: 978-1-80511-395-9 ISBN Digital (PDF): 978-1-80511-396-6 ISBN Digital eBook (EPUB): 978-1-80511-397-3 ISBN HTML: 978-1-80511-398-0

DOI: 10.11647/OBP.0425

Cover image: Golden Gate Bridge, San Francisco, California. Photo by Tj Kolesnik, https://unsplash.com/photos/silhouette-photography-of-mountain-beside-suspensionbridgeduring-golden-hour-Vnz7o7LeuDs.

Cover design: Jeevanjot Kaur Nagpal

In Chapter 3, I argued that to believe something is to take it as given in all relevant inquiries, and in at least one possible inquiry. I explained what it was to take something as given in terms of how one answers conditional and unconditional questions. In this chapter I'm going to argue that whatever is known can be properly taken as given in all relevant inquiries, where a relevant inquiry is one that one either is or should be conducting. Since some things that are usually known cannot be properly taken as given in some inquiries, this implies that knowledge is sensitive to one's inquiries and hence to one's interests.

There is an easy argument for the conclusion of this chapter.

- 1. To believe something is to, inter alia, take it as given for all relevant inquiries.
- 2. Whatever is known is correctly believed.
- 3. So, whatever is known is correctly taken as given in all relevant inquiries.

I think this argument is basically sound, but both premises are controversial. Further, it isn't completely obvious that it is even valid. So I'm not going to rely on this argument. Rather, I'll argue more directly for the conclusion that whatever is known is correctly taken as given in all relevant inquiries. This will provide indirect evidence that the theory of belief in Chapter 3 was correct, since we can now take that theory of belief to be an explanation for the claim that whatever is known is correctly taken as given in all relevant inquiries, rather than as part of the motivation for it.

The argument here will be in two parts. First, I'll focus on practical inquiries, i.e., inquiries about what to do, and argue that what is known can be taken as given in all practical inquiries. Then I'll extend the

discussion to theoretical inquiries, and hence to inquiries in general. Finally, with the argument complete, I'll look at two possible objections to the argument. One objection is that it has implausible consequences about the role of logical reasoning in extending knowledge, and the other is that it leads to implausible results when a source provides both relevant and irrelevant information.

4.1 Ten Decision Commandments

A practical inquiry can often be represented by the kind of decision table that we use in decision theory courses.¹ Table 4.1, for instance, is a table for the problem faced by a person, call him Ragnar, choosing how to get to work.

Table 4.1	Ragnar's	trip	to	work
fuble i.i	ragian b	urp	w	

	Rain	Dry
Walk	0	5
Bus	3	4

If we tell the students that the probability of rain is 0.4, we expect them to figure out that the expected utility of walking is 3, while the expected utility of taking the bus is 3.6. Therefore, taking the bus is the better choice. And that's a little surprising, since it probably won't rain, and if it doesn't, it is better to walk. The key point is that walking is risky, and in this case expected utility theory suggests that the risk isn't worth taking.

Table 4.1 can serve two related philosophical purposes, which we can helpfully distinguish using terminology from Peter Railton (1984). The table can provide a *criterion of rightness* for Ragnar's actions. It is rational for him to take the bus because of the expected utility calculation. The table can do more than that though. In simple cases like this one, it can provide a *deliberation procedure*. Ragnar can, in theory and in simple cases, use a table like this to decide what to do. There are limits to when tables can be used in this way, and as I'll argue in Chapter 6, those limits

¹ This section and the next are loosely based on Weatherson (2012: §1.1).

end up suggesting limits to how often the tables even provide criteria of rightness. In simple cases though, the table isn't just something the theorist can use to understand Ragnar, it is something Ragnar himself can use to deliberate. This is especially true in cases where one of the options is dominated, either strictly or weakly, by another.² I've appealed to the fact that the tables can be deliberation procedures, and not just criteria of rightness, already, in the discussion of Sully and Mack in Section 3.4. There the focus was on how tables like these related to belief; here I want to relate them to knowledge.

There are (at least) ten ways in which Table 4.1 could misrepresent Ragnar's situation. To put the same point another way, there are (at least) ten ways in which it could correctly represent his situation. One way to think about the core project of this book is to say what it means for a table to correctly represent a decision situation in one of these ten respects. It is a little easier to think about the misrepresentations, so I'll start with them.

First, the numbers in the table might be wrong. The table says that, conditional on catching the bus, Ragnar is better off if it is dry than if it rains. Maybe that isn't true. The theory of well-being (Crisp, 2021) addresses, among other things, when the numbers in the cells of tables like this are accurate. That's a big topic, and not one I'm going to have anything to say about here.³

Second, the probabilities might be wrong. Maybe it isn't the case that the probability of rain is 0.4, and in fact it is 0.2. There is an enormous question here about what it even means for one to misrepresent the probabilities. Is the correct representation one that tracks objective

² An option is strictly dominated by another if it does worse than that option in every state. It is weakly dominated by another if it does worse than that option in some states, and never does better than it.

³ As well as questions about well-being, there are also questions here about what one should do in cases where the outcome is itself a kind of gamble. Imagine that the chooser is trying to decide whether to bet on a basketball game, and it is known how much money they will win or lose in the four states. The value to the chooser of those outcomes depends on any number of further things, like the rate of inflation in the near term, and the "position of wealth holders in the social system" (Keynes, 1937: 214) some years hence. Just how these uncertainties should be accounted for is a difficult question, especially for any theorist who deviates in any way from orthodox expected utility theory. I would like to have a better theory of how the account of decision making with deliberation costs, as discussed in Chapter 6, interacts with this question.

chances, or Ragnar's evidence, or Ragnar's beliefs, or something else, or some combination of these? One upside of focussing on dominance arguments is that these questions can be temporarily set aside.

The next four questions concern the rows, and here we have less philosophical work to draw on. Brian Hedden (2012) has a paper arguing that the options should all be decisions, rather than actions. So the first row should say "Ragnar decides to walk" rather than "Ragnar walks". This would be a fairly radical change from practice in decision theory, though one worth taking seriously. The more conservative option would be to link the rows to some or other philosophical theory of abilities (Maier, 2022). In some sense it seems right to say that there should be a row for all and only the actions that Ragnar is able to perform. The details are going to be tricky though. This book is focussed on the columns rather than the rows, but I want to briefly mention four important topics about the rows, which will constitute our third through sixth ways the table might misrepresent Ragnar's situation.

Third, the table might leave off an option that should be there. Perhaps Ragnar should, or at least should consider, driving to work. Or perhaps it should include the option of quitting his job immediately, and hence not going to work.

Fourth, the table might include an option that should not be there. If the bus route near Ragnar's house has just been cancelled, perhaps the table should not include a row for the bus.

Fifth, the table might have merged multiple options that should be separated. Perhaps it should have separate rows for walking with an umbrella, and walking without an umbrella. This differs from the third point, because it does not say that Ragnar should do (or consider) something wholly distinct from what is already there, but rather that it should separate out different ways of bringing about something that is considered.

Sixth, the table might have separated multiple options that should be merged. It's hard to see how Table 4.1 could have made this mistake, but if we had separate rows for walking while wearing a red shirt, and walking while wearing a blue shirt, it would be arguable that this is too fine a grain, and the right table would not distinguish these.

The final four questions concern the columns, and they mirror the four questions about the rows. These questions will be central to the narrative of this chapter, and of this whole book.

4. Knowledge

Seventh, the table might leave off a state that should be there. Perhaps Ragnar should consider the possibility that it will snow, or that there will be an ice storm. Taking the only two states to be rain and dry excludes those possibilities,⁴ and perhaps they should be included.

Eighth, the table might include a state that should not be there. If it is bucketing down as Ragnar is preparing to leave, including a state where it is dry might be a mistake.

Ninth, the table might have merged states that should be separated. Perhaps the column that simply says "Dry" should have been split into two: one being "Dry and Sunny", the other being "Dry and Cloudy".

Tenth, the table might include separate states that should be merged. It's unlikely that a two state table will do this, but if we had made the split suggested in the previous paragraph, one could easily argue that it was a mistake, and that Ragnar should have treated these as a single state.

That gives us ten ways that the table could go wrong. It's helpful to have them in a simple list.

- 1. The values could be wrong.
- 2. The probabilities could be wrong.
- 3. An option could be improperly excluded.
- 4. An option could be improperly included.
- 5. The options might be too coarse-grained.
- 6. The options might be too fine-grained.
- 7. A state could be improperly excluded.
- 8. A state could be improperly included.
- 9. The states might be too coarse-grained.
- 10. The states might be too fine-grained.

For every one of these ten possible mistakes, there is a prior philosophical question about what it means for the table to have made, or not made, that mistake. Every one of those ten questions is, at least to my mind, incredibly philosophically important. Even someone who thought, like Herbert Foxwell, that books should only be written for "grave cause"

⁴ As noted back in Section 3.4, I'm using 'possibilities' here in the sense described by Humberstone (1981).

(Keynes, 1936: 599), should concede that a clear answer to any one of the ten would be sufficient grounds to warrant a scholarly monograph.

This book is primarily concerned with the seventh, though the argument touches to some extent on the eighth as well. It is proper to exclude a possibility from the table if the chooser knows that possibility does not obtain. If that conditional could be turned into a biconditional, we'd have an answer to the eighth question, too, but that is a more delicate question.⁵ In any case, the conditional will be enough.

4.2 Knowing Where the Ice Cream Goes

The aim of this section is to argue for the following principle.

Knowledge Allows Exclusion (KAE)

If a chooser knows that a possibility does not obtain, then it is permissible to use a decision table where that possibility is excluded, i.e., is incompatible with the possibilities in each of the columns.

Knowledge Allows Exclusion is Jessica Brown's principle K-Suff applied to practical decision making using tables. That's a fairly central case for K-Suff, so if KAE is true, then it seems plausible that K-Suff will be true too. I'll come back to the more general case for K-Suff in later sections, though; here the focus is KAE. I'm going to build up to KAE in stages; first I'm going to talk about ice cream.

The contemporary theory of duopoly starts with Harold Hotelling's paper "Stability in Competition" (1929). Hotelling describes how a duopoly that does not maximise consumer welfare can be stable if the two parties have the ability to differentiate their product along one dimension. Surprisingly, the equilibrium is that they do not in fact take advantage of this ability, and instead provide the very same product. Hotelling's observation is that if both parties could differentiate, neither party has the incentive they would normally have to reduce prices to the point where consumer surplus is maximised. Hotelling is interested in possible equilibria, and he doesn't focus on how the parties might calculate the equilibria. (The impression one gets from the paper is that it will involve a good chunk of trial-and-error.) Subsequent work revealed that it turns

⁵ Back in Section 2.7 I said I was staying neutral on that question, and I'm not changing that position here.

4. Knowledge

out that in some duopoly situations, not much is needed to get to the equilibrium; just iterated deletion of dominated strategies.

Here is the standard way Hotelling's model is introduced in textbooks.⁶ Imagine that two ice cream trucks have to choose (simultaneously) where they will be located on a beach. The beach has seven locations, numbered 1 to 5. The distance between location m and location n is |m - n|. Assume for simplicity that the price of ice cream is fixed, the trucks just compete on location. There are two beach-goers at each of locations 1 to 5, so 10 in total. Each beach-goer will buy an ice cream from the nearest truck. If two trucks are equidistant from a location, the two people there will head off in either direction, one buying from each truck. Question: Where should the two trucks go, assuming that it is common knowledge that each truck owner is rational, and simply wants to maximise their own sales?

This puzzle can be solved using just the idea that strictly dominated strategies can be iteratively deleted. Table 4.2 shows how many sales each truck will make for each choice of location. The choice of the first truck determines which row of the table we're in, the choice of the second truck determines which column of the table we're in, and the resulting cell lists first the sales of the first truck, then the sales of the second truck. (So we'll call the first truck Row, and the second truck Column.)

	Table	e 4.2 Payouts in	the Hotelling g	game.	
	1	2	3	4	5
1	5,5	2,8	3,7	4,6	5,5
2	8,2	5,5	4,6	5,5	6,4
3	7,3	6,4	5,5	6,4	7,3
4	6,4	5,5	4,6	5,5	8,2
5	5,5	4,6	3,7	2,8	5 <i>,</i> 5

6 This particular example isn't in Hotelling, but it is in so many textbooks that I haven't been able to find out where it was first introduced. It differs from his examples in that the parties do not have the capacity to compete on price.

Assume that it is common knowledge, in the sense of David Lewis (1969), that Table 4.2 is the payout table, and that each player will not make choices that are strictly dominated. That is, for each *n*, the proposition we get by having *n* iterations of *each player knows* in front of *this is the game table, and each player is rational*, is true. Then the theorist, and each player, can reason as follows.

Row's option 1 is strictly dominated by option 2; option 2 gets 1 more sale in three possible states, and 3 more sales in the other two, so it should be excluded. The same goes for option 5, which is strictly dominated by option 4. Since the game is symmetric, the same goes for Column's options 1 and 5. By the common knowledge assumption, this means we can delete those rows, and columns, from the table. The result is Table 4.3.

Table 4.3 The Hotelling game after one iteration.

	2	3	4
2	5,5	4,6	5,5
3	6,4	5,5	6,4
4	5,5	4,6	5,5

For both players, option 3 dominates the other two options, so it will be chosen. Moreover, the reasoning here generalises. If there are seven options to start with, we need to do two rounds of deleting dominated options to get the players to the middle of the beach. If there are nine options to start with, we need to do three rounds of deletion. In general, if there are 2k+1 options, we get the players to the middle of the beach after *k* rounds of deletion. Since common knowledge licences all these iterations, the players will always end up in the middle of the beach if there are an odd number of options.

At this point you might be worried for two reasons. Practically, this seems like it proves too much. Contra the conclusion of Hotelling's paper, it's not true that shoes, churches, and cider mills are as homogenous as this argument would suggest. Theoretically, there are plenty of reasons to be worried about common knowledge as Lewis understood it. Harvey Lederman (2018) shows that assuming common knowledge, in Lewis's sense, of dominance avoidance leads to paradoxes. Let's see whether we can get by with less.

Assume that it is not common knowledge, but merely mutual knowledge that the payout table is as in Table 4.2, and the players do not take dominated options. That is, each player knows both those things. That is all we'll assume. Since knowledge is factive, we can still rule out the extreme options, i.e., 1 and 5. Given that each player knows the other will not take dominated options, each player knows that it is only options 2 through 4 that are relevant. So given just the mutual knowledge assumption, we can show that from each player's perspective, they are playing the game depicted in Table 4.3. In that game, option 3 is strictly dominant. So this assumption is enough to get us back to the middle of the beach. Note, however, that this reasoning does not generalise. Given merely mutual knowledge of non-dominance, we can show that neither player will take options 1 or 2, or the second-last or last options, but we can't show any more than that. So in the seven-option game, we can only show that they will both end up somewhere between options 3 and 5. In the games with much larger numbers of options, we can't show much at all. That seems both empirically and theoretically more plausible.

The argument of the last paragraph is meant to serve two distinct, but related, philosophical purposes.⁷ First, it is meant to show that we theorists can deduce what the players will in fact do, given their evidence, and the assumptions about rationality. Second, it is meant to show that it would be rational for the players themselves to get to that conclusion via just that reasoning. It is important, in general, to distinguish between what is entailed by some assumptions, and what can be reasonably inferred from those assumptions (Harman, 1986). In this case, though, I want to claim that the reasoning I've set out in that paragraph plays both roles. As theorists, we can tell that the players will not play either the extreme, or the next to extreme, option, and no more. The players themselves will not go to any of those 4 spots, given our assumptions, but we can't know more about their actions without more knowledge of their mental states.

⁷ I'm indebted here to conversations with Eric Swanson.

But wait a minute! Without KAE, the last two paragraphs consist of one fallacious step after another. The player knows that the other player will not play an extreme option. Also, they know that if the extreme options are excluded, option 2 is strictly dominated. Without KAE, it doesn't follow that they can simply delete the extreme options. To delete an option just is to exclude it from the table. Without KAE, the fact that the player knows an option doesn't obtain isn't a sufficient reason to make this deletion. Since it is, in practice, a sufficient reason, it follows that KAE is true. Or, at least, that a restricted version of KAE applied to this case is true. Since the case seems arbitrary, it follows that KAE is true in general.

That's my primary argument for KAE. In general, it is reasonable to do as many rounds of deletion of dominated strategies as we have iterations of mutual knowledge of rationality and the structure of the game table. That is, it is reasonable for the theorist to do exactly as many rounds of deletion as there are iterations of mutual knowledge of rationality among the players. Without KAE, that match up isn't guaranteed, so KAE must be true.

4.3 Other Answers

If KAE is false, what should go in its place? What could be the state which does allow exclusion?

4.3.1 None of the Above

One might object to the presupposition of that question. Maybe exclusion is never allowed. Perhaps every table should partition the possibility space. In any table, the last state should be *None of the above*, so (assuming classical logic) it must always be true that some state in the table obtains.

If one is not completely convinced that classical logic is correct, this move won't seem particularly appealing. I suspect, however, that most readers are completely convinced that classical logic is correct, so I won't investigate that line. Instead, I'll look at two more pressing objections to the idea that decision tables should always have a *None of the above* option.

First, in many cases there is no sensible way to determine the probabilities or utilities that would go in this column. Imagine that I'm making a decision whose consequences are sensitive to which team wins the next Super Bowl. (Perhaps I'm planning a giant Super Bowl party, or I'm setting the odds for season-long bets at a sportsbook.) I work out the probabilities that each of the 32 teams in the National Football League (NFL) will win this year, and what the consequences of my various options would be in each case. If it's never permissible to exclude states from a decision table, if decision tables always have to be logically complete, I need a 33rd state: that none of these teams win. But how could that be? Maybe the League might be cancelled? Maybe a new team could be introduced mid-season and could win? There is not really a sensible way to even assign probabilities to these options. Worse still, there is no way to assign utilities to actions given that state. The expected return of an action given this state will depend on the probabilities of the different ways it could come about. The error bars on those probabilities are bigger than the probabilities themselves. There is simply no sensible value to put in the cell as the value of the pair Schedule a large Super Bowl party in Las Vegas and None of these 32 teams win the league. If that state comes about because the Super Bowl is cancelled, it's terrible. If it comes about because a new team gets added, that would create so much interest that it would be great. If I don't have any way of figuring out the relative probabilities of these events, I have no idea what the expected value is. So this approach makes decision tables useless.

Second, one should only be unwilling to exclude states from decision tables if one is so sceptical that one is unwilling to take any contingent proposition to be evidence. After all, taking something to be evidence involves excluding possibilities where it doesn't obtain from one's reasoning. If one doesn't take anything to be evidence, then it is unclear how one's probabilities can update. It can't be by regular conditionalisation. It could be by Jeffrey conditionalization (Hájek 2011: §7.2), if one thought that somehow it was impossible to ever learn that p, but sometimes possible to learn what p's probability is. Personally, I've never had a learning experience that told me the precise probability of some proposition without learning for sure some other proposition. I have never seen reason to think anyone else has either.

This is a quite general point about interest-relative epistemology, and one that will keep coming up in different ways throughout the book. If one wants to do without knowledge, and just use probabilities (or credences), one owes us a story of how those probabilities change. The best stories about how probabilities change all involve some kind of interest-relativity.

4.3.2 Evidence

These considerations suggest a different answer to this exclusion problem; perhaps the decision maker can exclude p iff p is part of their evidence. Call this view EAE, for Evidence Allows Exclusion.

It isn't obvious that this is an alternative to KAE. If evidence and knowledge are co-extensive, as Timothy Williamson (2000) argued, it will not be. Since I'm going to argue in Chapter 9 that Williamson is wrong about this, I'm committed to EAE and KAE being distinct. So I need an argument against EAE.

My argument will be by cases. That p is part of one's evidence either entails that one knows p or does not. Either way, EAE doesn't pose a problem for my overall argument.

If it does, then whether EAE or KAE is true won't matter for the overall argument. I'm going to argue that some propositions that are known in typical situations might not be properly excluded if one's interests change. That will imply interest-relativity given KAE, but it will also imply interest-relativity given EAE plus the thesis that evidence entails knowledge.

If evidence doesn't entail knowledge, then EAE is implausible. If evidence isn't strong enough to let the decision maker know that propositions inconsistent with it are false, it surely isn't strong enough to let the decision maker know they can ignore propositions inconsistent with it.

The view I'll defend in Chapter 9 is that evidence does entail knowledge. There is a really simple argument for this view. One way to know that p is by properly deducing p from one's evidence. The deduction p, therefore p can be properly carried out. So one can know anything in one's evidence. I'm not relying on this argument here, and instead on the point that if evidence doesn't suffice for knowledge, it surely doesn't suffice for exclusion.

The same considerations show that CAE, the view that Certainty Allows Exclusion, doesn't threaten the larger argument for interest-relativity. Either certainty entails knowledge or it doesn't. If it does, then CAE can be used in place of KAE below to derive interest-relativity. If it does not, and this might happen if certainty just is subjective certainty, then it is implausible that it suffices for proper exclusion.

4.3.3 Sufficiently High Probability

Perhaps one can exclude those propositions whose falsity is sufficiently high that treating them as definitely false doesn't make a difference to the decision one makes. Call this view PAE, for sufficiently high Probability Allows Exclusion.

The first thing to note is that if this is to be plausible, the notion of sufficiency here must be interest-relative. It's often fine to ignore propositions that have a 1 in 500 chance of being true. When planning what to do on a fine sunny day with a clear weather forecast, I simply ignore the chance that there will be a passing shower, even though that still has a 1 in 500 chance. On the other hand, it's absurd to ignore 1 in 500 chances when deciding what insurance to buy. About 1 house in 500 has a fire in a given year; that's not a conclusive reason to skip fire insurance for the year.

Second, as stated, this view has the odd consequence that decision makers can ignore situations that actually obtain. This doesn't seem very plausible. At least, it would be very odd to have a textbook representation of a decision problem where the actual world wasn't in one of the columns. So probably the best way to interpret PAE is as saying that falsehoods can be excluded iff they have sufficiently high probability.

Third, once one does that, PAE starts to look suspiciously like a form of KAE. In particular, it looks like the view I'll call IRT-CP in Chapter 6. That means (a) that it isn't obviously an alternative to KAE, and (b) the objections to IRT-CP are also objections to it. Since I'll go over those objections in detail in Section 8.2 and Section 8.3, I won't double them up here, but assume that they work against PAE.

4.3.4 Wrapping Up

I've argued that the states we can exclude from a decision table are the states that the agent knows not to obtain. The argument is largely by elimination. One might object that I haven't excluded *all* alternatives.

We could keep going asking whether one can exclude all and only those things that are justifiably believed to be false, or which are known to be known, or any number of other alternatives.

At this point, it is natural to object that some alternatives are too complicated to warrant much confidence. What we can properly take for granted in decision making is a very important fact about our doxastic states. If one is sympathetic to a broadly functionalist picture of mind, it might be the most important fact. If so, it isn't surprising that the most common form of appraisal of doxastic states, that they are knowledge, is the norm for appropriate exclusion. It would be very surprising if something considerably more complicated was the correct norm instead.

That's hardly a conclusive argument, but it seems like a good enough one to leave off the survey here, and return to the main narrative of asking what follows if Knowledge Allows Exclusion.

4.4 From KAE to Interest-Relativity

If KAE is true then there is a simple argument that Anisa loses knowledge when playing the Red-Blue game. Table 4.4 would be a bad table for Anisa to use when deciding what to do.

Table 4.4 What the Red-Blue game looks like if Anisa assumes that the Battle of Agincourt was in 1415.

	2+2=4	$2+2 \neq 4$
Red-True	\$50	0
Red-False	0	\$50
Blue-True	\$50	\$50
Blue-False	0	0

If she used that table, then it would look like Blue-True is the weakly dominant option. That would mean that Blue-True is at least a rational choice, and perhaps the rational choice. Since Blue-True is not a rational choice, this table must be wrong. If Anisa knows that the Battle of Agincourt was in 1415, and knowledge structures decision tables, then

everything on this table is correct. So Anisa does not know that the Battle of Agincourt was in 1415. Since she does know this when not playing the game, her knowledge is interest-relative.

4.5 Theoretical Knowledge

Knowledge structures proper practical deliberation. Because what can be taken as a structural assumption varies across different instances of practical reasoning, knowledge is sensitive to the interests of the inquirer. But this isn't the only way in which knowledge is sensitive to interests. It is also sensitive to which purely theoretical questions the inquirer is taking an interest in.

I've already mentioned one way in which this has to be true. One kind of theoretical question is *What should I do in this kind of situation?* If actually being in that kind of situation and having to decide what to do affects what one knows, then thinking abstractly about it should affect what one knows as well.

This kind of comparison, between practical deliberation about what to do and theoretical deliberation about what one should do in just that situation, suggests a few things. It suggests that if practical interests affect knowledge, then so do theoretical interests. It also suggests that they should do so in more or less the same way. So it would be good to have a story that assigns to knowledge the role of structuring theoretical deliberation, in just the way that it structures practical deliberation. That's more or less the story I'm going to tell, though there are some complications along the way.

The story I like starts with an observation by Pamela Hieronymi.

A reason, I would insist, is an item in (actual or possible) reasoning. Reasoning is (actual or possible) thought directed at some question or conclusion. Thus, reasons must relate, in the first instance, not to states of mind but to questions or conclusions. (Hieronymi, 2013: 115–116)

So to a first approximation the inquirer knows that p only if they can properly use p as a reason in "thought directed at the question" they are considering. That is, they can use p as a step in this reasoning. This way of putting things connects Hieronymi's view of reasons to the idea present in both John Hawthorne and Jason Stanley (2008) and Jeremy Fantl and Matthew McGrath (2009) that things known are reasons. While I'm going to spend the rest of this section quibbling about whether this is quite right, it's a good first step.

It's enough to get us a fairly strong, but also fairly natural, kind of interest-relativity. In normal circumstances, Anisa knows that the Battle of Agincourt was in 1415. Now imagine not that she's playing the Red-Blue game, but thinking about how to play it. And she wonders what to do if the red sentence says that two plus two is four, and the blue sentence says that the Battle of Agincourt was in 1415. It would be a mistake for her to reason as follows: well, the Battle of Agincourt was in 1415, so playing Blue-True will get me \$50, and nothing will get me more than \$50, so I should play Blue-True. The mistake is the first step; she just can't take this for granted in this very context.

This is a very obscure kind of question to wonder about, but there are more natural questions that lead to the same kind of result. Imagine that the day after reading the book, but before playing any weird game, Anisa starts wondering how likely it is that the book was correct. History books do make mistakes, and she wants to estimate how likely it is that this was a mistake. Again, it would be an error to reason as follows: well, the Battle of Agincourt was in 1415, and that's what the book says, so the book is certainly correct. Again, the problem is the first step; she just can't take this for granted in this very context.

But it's not as if she can only take things for granted that are certain in that context. If that were true, she couldn't even start an inquiry into how likely it is the book got this wrong. She has to assume certain things are beyond the scope of present inquiry. She should not question that the book says that the battle occurred in 1415, or that there was a Battle of Agincourt, or that it is a widely written about (but also widely mythologised) battle, or that 1415 is before the invention of the moveable type printing press and so records from 1415 might be less reliable, and so on. None of these things are things that she knows with Cartesian certainty. Indeed, some of them are probably all-things-considered less likely than that the Battle of Agincourt was in 1415.⁸ So it's not like there is some threshold of likelihood, or of evidential support, and inquiring into the likelihood of this statement implies that one can take for granted all and only things that clear this threshold. Rather, individual inquiries

⁸ When I was editing this book I realised I wasn't sure when the moveable type printing press was invented, and had to double check it was after 1415.

have their own logic, their own rules about what can and can't be taken for granted.

There is an interesting analogy here with the rules of evidence in criminal trials. Whether some facts can be admitted at a trial depends in part on what the trial is. For example, some jurisdictions allow evidence obtained in a search that illegally violated X's rights to be used in a trial of Y, though it could not be used when X was on trial. The picture I have of knowledge is similar; what one knows is what one can use in inquiry, and what one can use changes depending on the question under discussion. I'll have much more to say about this in Chapter 5.

So the starting point is that what's known is what can be used. What I'm going to ultimately defend is a much more restricted thesis. Using what is known provides immunity from a particular criticism: that your starting point might not be true. I'm going to say a little bit about why this immunity claim is correct, and then say much more about why I prefer this way of talking about the role of knowledge in reasoning.

When one says that it is good to use what one knows in reasoning, there are two natural ways to interpret this. One is that using what one knows is all-things-considered good unless there is some independent reason to the contrary. The other is to say that there is a kind of badness in reasoning one avoids if one uses what one knows. I'm going to be defending the second kind of reading. That's what I mean by saying that using what one knows provides immunity from a certain kind of criticism. The alternative requires that we can specify all the ways in which one might go wrong while using what one knows—those are the "independent reasons to the contrary". I don't think that's something we're now in a position to do.

The justification for the immunity claim is quite straightforward. It's incoherent to say of someone that they know that p, but they shouldn't have used p in reasoning because it might be false. That's Mooreparadoxical, if not outright contradictory. If it is incoherent to say A, and X shouldn't have done B because C, then A is a good reply to the criticism of X that she shouldn't have done B because C. So knowing that p is a good reply to the criticism that one shouldn't have used p in reasoning because it might be false.

Can we say something stronger? Can we say that knowing that p immunises the reasoner from all criticisms? Surely not; using irrelevant facts in inquiry is a legitimate criticism, even if the facts are known

(Ichikawa, 2012). Is there a true claim that's a bit more qualified, but still stronger than the immunity claim that I make?

One possibility would be to say that reasoning that starts with what is known is immune from all criticisms except those on a specified list. What might be on the list? I've already mentioned one thing—using irrelevant facts. Another thing might be that the reasoning itself is irrelevant to what one should be doing. If there is a drowning child in front of me, and I start idly musing about what the smallest prime greater than a million might be, I can be criticised for that reasoning. That criticism can be sustained even if my mathematical reasoning is impeccable, and I get the correct answer.⁹

Some facts are irrelevant to an inquiry. Others are relevant, but not part of the best path to resolving the inquiry. This can be grounds for criticism as well. It's in some cases a mild criticism. If one follows an obvious path to solving a problem, when there is an alternative quicker way to solving the problem using a clever trick, it isn't much of a complaint to say that the reasoning wasn't maximally efficient. There are many quicker proofs of a lot of things Euclid proved, but this hardly detracts from the greatness of Euclid's work. And, interestingly for what is to follow, using an inefficient means of inquiry does not prevent the inquiry ending in knowledge. After all, Euclid knew a lot of geometry, even though he rarely had maximally efficient proofs. There is a general lesson here—the fact that an inquirer was imperfect isn't in itself a reason to deny that they end up with knowledge.

Inefficiency in inquiry is often not a big deal; other mistakes in inquiry are more serious. Sometimes the premises do not support the conclusion. It's notoriously hard to say what is meant by support here. It seems to have some rough relationship to logical entailment, but it's hard to say more than that. Sometimes premises support a conclusion they do not entail—that's what happens in all inductive inquiry. Sometimes premises do not support a conclusion they do entail. If I reason, "3 is the first odd prime greater than 0, so 1,000,003 is the first odd prime greater than 1,000,000, and there are no even primes greater than 2, so 1,000,003 is the first prime greater than 1,000,000. But the

⁹ As it turns out, that's 1,000,003.

premise, that 3 is the first odd prime greater than 0, entails the next step. It just fails to support it, in the relevant sense.

Maybe now we might suspect we've got enough criticisms on the table. Is there anything wrong about an inquiry where the following criteria are met?

- It is worthwhile to conduct the inquiry.
- It is sensible, and efficient enough, to choose these particular starting points.
- The starting points are all things that are known to be true.
- Every step after the starting point is supported by the steps immediately preceding it.

An inquiry with these features looks pretty good. If there is really nothing to complain about in such an inquiry, then the following is true. An inquirer who starts an inquiry with what they know is immune from all criticisms except perhaps (a) that they shouldn't be conducting this inquiry at all, (b) that their starting points are irrelevant (or perhaps inefficient) for reaching their conclusion, or (c) that their later steps are not supported by their earlier steps. While those are fairly non-trivial exception clauses, that's still a fairly strong claim about the role of knowledge in inquiry.

Unfortunately, there are puzzle cases that suggest that even an inquiry with those four features may be flawed. I'll just mention two such cases here. The point of these cases is that they suggest inquiry can be flawed in ever so many ways, and we should not be confident about putting together a complete list of the ways inquiry can go wrong.

First, there might be moral constraints on inquiry. Consider the following example, drawn from Rima Basu and Mark Schroeder (2019). Casey is at a fancy fundraising party, where the guests and the wait staff are all wearing suits. The person next to Casey is Black, and Casey reasons as follows.

- 1. Almost all the Black people here are on the wait staff.
- 2. The person next to me is Black.
- 3. So, the person next to me is on the wait staff.

That's not valid, but one might argue that it's a rational inductive inference. Alternatively, we can consider the case where Casey explicitly

concludes that the person next to them is probably Black. We can imagine that all of the following things are true. It is reasonable for Casey to think about whether the person in question is on the wait staff; it matters for the reasonable practical purpose of getting a drink. The wait staff are not wearing distinctive clothes, so seeing what observational characteristics correlate with being on the wait staff is a reasonable approach to that inquiry. Casey knows that the premises of the inquiry are true, and the premises support the conclusion of the inquiry.

And yet, it seems something goes badly wrong if Casey reasons this way. If the conclusion is false, it doesn't seem like mere inductive bad luck. Arguably, there is a moral prohibition on reasoning in this way. Furthermore, this moral prohibition plausibly prevents Casey's reasoning from providing knowledge.

Now one might well question just about every step of the last two paragraphs. It's one thing to regret the lack of signals from attire as to who is on the wait staff; it's another thing to jump to using skin colour as the best proxy. Given how many other things Casey can see about this person (such as how they are moving, what they are carrying, how they are engaging with others), it isn't clear that the premises support the conclusion, even inductively.

Even if all those things are not true, it might be that Casey can get knowledge this way; the inquiry might be morally wrong without having any epistemic flaws that prevent it generating knowledge. Other examples of morally problematic inquiry suggest that there is no simple connection between an inquiry being morally bad, and it not generating knowledge. Many inquiries are morally problematic because they involve, or even constitute, privacy violations. But that doesn't mean the privacy violator doesn't come to know things about their victim. Indeed, part of the wrongness of the privacy violation is that they do come to know things about their victim.

Still, Casey can be criticised for inquiring in this way, even if the criticism does not imply that the inquiry produced no knowledge. That suggests that there are possible criticisms of inquiries that satisfy the four bullet points listed earlier.

Another source of trouble comes from holistic constraints on reasoning. What I have in mind here are rules that allow for a natural resolution of the puzzles of "transmission failure" that Crispin Wright (2002) discusses. Start with one of Wright's examples. Ada is walking by a park with a football pitch. It clearly isn't just a practice; the players are in uniforms and occupying familiar positions on the pitch, there is a referee and a crowd, and so on. One of the players kicks the ball into the net, the referee points to the centre of the ground, and half the players and crowd celebrate. After this happens, Ada reasons as follows.

- 1. The ball was kicked into the net, and no foul or violation was called.
- 2. So, a goal was scored.
- 3. So, a football match is being played, as opposed to, e.g., an ersatz match for the purposes of filming a movie.

As Wright points out, there is something wrong with the step from 2 to 3 here. As he also points out, it isn't trivial to say just what it is that's wrong. After all, 2 entails 3, and Ada knows that 2 entails 3. But it seems wrong to make just this *inference*.

Here's one natural suggestion about what's wrong.¹⁰ It's too simple to be the full story, but it's a start. The transition Ada makes from 1 to 2 presupposes 3, and 1 is her only evidence for 2. When those two conditions are met, it is wrong to infer from 2 to 3. More generally, there is something wrong with inferring a conclusion from an intermediate step in reasoning if that conclusion must be presupposed in order to even reach that intermediate step.

This is too rough as it stands to be a full theory of what is going on in cases like Ada's, but the details aren't important at this point. What is important is that there might be some kind of holistic constraint on reasoning. In some sense, Ada goes wrong in taking 2 for granted when she infers 3. This doesn't intuitively undermine her claim to know both 2 and 3.

One important commonality between the last two cases, the moral encroachment and the transmission failure cases, is that the reasoning is not subject to the following kind of criticism. The speaker can't be criticised for taking as a premise something that might be false. Maybe there is something wrong with inferring something is probably true of an individual because it is true of most people in the group of which

¹⁰ This is far from an original suggestion. See Weisberg (2010) for discussion of it, and of related proposals, and for more discussion of the literature on Wright's examples.

the individual is part. But this restriction applies to the inference, not to the premises. We wouldn't say to the person who made this inference, "You shouldn't reason like that; it might not be true that most people in the group have this feature." If we did say that, they would have an easy reply. If Ada does do the problematic reasoning, it would be wrong to reply to her "You shouldn't reason like this; it might not have been a goal." She could simply, and correctly, say that it quite clearly was a goal.

This is the key to the correct rule linking knowledge and reasoning. If the inquirer uses as a step in reasoning something that she knows to be true, then she is immune to a certain kind of criticism. She is immune to the criticism that the premise she used might not be true.

I started this section by saying that such a reasoner is immune to all criticism, before trying to work out exceptions to that principle. So an exception needed to be included to allow that the reasoner might be criticised for using an irrelevant reason. The hope was that eventually a full list of such exceptions could be found. This project turned out to be wildly optimistic. I don't know that we need to include further exceptions to handle the moral encroachment or transmission failure cases. But I also don't know that we don't need to include extra exceptions. And I have no idea, and no idea how to find out, whether we need yet more exceptions.

Rather than say knowledge provides immunity to criticism except in these cases, and then try to fill out the list of cases, it's better to say that knowledge provides a particular kind of immunity. If the reasoner knows that the premise they use is true, they can't be criticised on the grounds that it might be false. This isn't a trivial claim. There were several examples involving Anisa where she could be criticised for using a premise that might be false. All of those seemed like legitimate criticisms even though the premise was one she knew before starting the inquiry. That criticism does not seem appropriate in the moral encroachment case, or the transmission failure case, or other cases like them that may be discovered.

I am assuming here that there is no trivial connection between *It might be that not-p*, and *The inquirer does not know that p*. If these claims express the same thing, at least in the particular context of evaluating the inquirer, then it would be trivial to say that knowledge provides immunity to criticism on the grounds that one's premises might not

be true. The recent literature on epistemic modals, however, does not inspire confidence that any such trivial connection exists.¹¹ So this immunity seems like a non-trivial claim.

So the key principle I'll be working with is that one cannot be criticised for using what one knows in an inquiry on the grounds that one is using what might be false. That's a bit of a mouthful, so sometimes I'll simply say that one can rationally take for granted what one knows. I'll have a lot more to say about this principle in the rest of this book, especially in Chapter 9.

I'll spend the rest of this chapter talking about how this principle relates to the idea that knowledge is closed under competent deduction. There are interesting examples that seem to show that the principle leads to several distinct kinds of violations of that principle. I'll argue that this is not right, and for any plausible closure principle, adding the idea that one can take for granted what one knows does not yield a new objection to that principle.

The principle as stated is a little ambiguous, and to defend it I need to resolve that ambiguity. Surprisingly, I need to resolve it by taking the logically stronger disambiguation. Normally if a principle is ambiguous, and might lead to problems, the trick is to insist on the weaker reading. That's not what's about to happen.

When I say that an inquirer can rationally take for granted the things they know, this should be understood collectively. That's to say, I endorse the collective, and not (merely) the individual, version of the immunity to criticism principles stated here.

Take for Granted (Individual)

If an inquirer knows some things, then each of those things are such that they can take that thing for granted in conducting the inquiry.

Take for Granted (Collective)

If an inquirer knows some things, then they can take all of those things for granted in conducting the inquiry.

¹¹ See Holliday and Mandelkern (2024) for a survey of how differently the two claims behave in embeddings and inferences, and a radical claim about how to best account for those differences.

I'll come back to the difference between these principles, and why I need to endorse the collective version, in Section 4.6.2. Until then I'll be talking about single pieces of knowledge at a time.

4.6 Knowledge and Closure

Here are two very plausible principles about knowledge, both taken from John Hawthorne (2005).

Single Premise Closure

If one knows p and competently deduces q from p, thereby coming to believe q, while retaining one's knowledge that p, one comes to know that q. (Hawthorne, 2005: 43)

Multiple Premise Closure

If one knows some premises and competently deduces q from those premises, thereby coming to believe q, while retaining one's knowledge of those premises throughout, one comes to know that q. (Hawthorne, 2005: 43)

Hawthorne endorses the first of these, but has reservations about the second for reasons related to the preface paradox. I'm similarly going to endorse the first and have reservations about the second. But my reasons don't have anything to do with the preface paradox. I argued in "Can We Do without Pragmatic Encroachment?" (Weatherson, 2005a) that concerns about the preface paradox are over-rated, and I think those arguments still hold up. But I have a slightly different qualification than Hawthorne does to Multiple Premise Closure, and I will discuss that more in Section 4.6.2.

It is not trivial to prove that my version of IRT satisfies these closure conditions. One reason for this is that I have not stated a sufficient condition for knowledge. All that I have said is that knowledge is incompatible with a certain kind of caution. So in principle I cannot show that if some conditions obtain then someone knows something. What I can show is that introducing new conditions linking knowledge with relevant questions does not introduce new violations of the closure conditions.

4.6.1 Single Premise Closure

But it turns out that even showing this is not completely trivial. Imagine yet another version of the Red-Blue game.¹² I'll call the player for this version of the game Margit, and I'll assume at various points that she acts rationally.

In this game, both of the sentences are claims about history that are well supported without being certain. And both of them are supported in the very same way. It turns out to be a little distracting to use concrete examples in this case, so just call the claims *A* and *B*. Imagine that Margit read both of these claims in the same reliable but not infallible history book, and she knows the book is reliable but not infallible, and she aims to maximise her expected returns. Then all four of the following things are true about the game.

- 1. Unconditionally, Margit is indifferent between playing Red-True and playing Blue-True.
- 2. Conditional on *A*, Margit prefers Red-True to Blue-True, because Red-True will certainly return \$50 while Blue-True is not completely certain to win the money.
- 3. Conditional on *B*, Margit prefers Blue-True to Red-True, because Blue-True will certainly return \$50 while Red-True is not completely certain to win the money.
- 4. Conditional on $A \land B$, Margit is back to being indifferent between playing Red-True and playing Blue-True.

From 1, 2, and 3, it follows in my version of IRT that Margit does not know either *A* or *B*. After all, conditionalising on either one of them changes her answer to a relevant question. The question being, *Which option maximises my expected returns*?, where this is understood as a mention-all question.

Now see what happens at point 4. Conditionalising on $A \land B$ does not change the answer to that question. So, assuming there is no other reason that Margit does not know $A \land B$, arguably she does know $A \land B$. That would be absurd; how could she know a conjunction without knowing either conjunct?

¹² This game will resemble the examples that Zweber (2016) and Anderson and Hawthorne (2019b) use to raise doubts about whether pragmatic theories like mine really do endorse Single Premise Closure.

Here is how I used to answer this question. Define a technical notion of interest. Say that a person is interested in a conditional question If p, Q? if they are interested, in the ordinary sense, in both the true-false question p? and they are interested in the question Q?. If conditionalising on a proposition changes (or should change) their answer to any question in which they are interested in this technical sense, then they don't know that proposition. This solves the problem because conditionalising on $A \wedge B$ does change their answer to the question If A, which option maximises expected returns? on its mention-some reading. So even though 4 is correct, this does pose a problem for closure.

This was not a great solution for two reasons. One is that it seems extremely artificial to say that someone is interested in these conditional questions that they have never even formulated. Another is that it is hard to motivate why we should care that conditionalisation changes (or should change) one's answers to these artificial questions.

There was something right about the answer I used to give. It is that we should not just look at whether conditionalisation changes the answers a person gives to questions they are interested in. We should also look at whether it changes things 'under the hood'; whether it changes how they get to that answer. The idea of my old theory was that looking at these artificial questions was a way to indirectly look under the hood. What I got wrong was trying to find some other question whose answer changed when and only when what was under the hood changed. I should have just looked under the hood.

So let's look again at the two questions that are relevant. This time, don't think about what answer Margit gives, but about how she gets to that answer.

- 1. Which option maximises expected returns?
- 2. If $A \wedge B$, which option maximises expected returns?

In the most natural interpretation of what Margit does, there will be a step in her answer to 5 that has no parallel in her answer to 6.

She will note, and rely on, the fact that she has equally good evidence for *A* as for *B*. That is why each option is equally good by her lights. The equality of evidence really matters. If she had read that *A* in three books, but only one of those books added that *B*, then the two options would

4. Knowledge

not have the same expected returns. She should check that nothing like this is going on; that the evidence really is equally balanced.

But nothing like this happens in answering 6. In that case, $A \wedge B$ is stipulated to be given. So there is no question about how good the evidence for either is. When answering a question about what to do if a condition obtains, we don't ask how good the evidence for the condition is. We just assume that it holds. So in answering 6, there is no step that acknowledges the equality of the evidence for both *A* and *B*.

So in fact Margit does not answer the two questions the same way. She ends up with the same conclusion, but she gets there by a different means. And that is enough, I say, to make it a different answer. If she knew $A \land B$ she could follow exactly the same steps in answering 5 and 6, but she cannot.

What should we say if she does follow the same steps? If this is irrational, nothing changes, since what matters for knowledge is which questions should be answered the same way, not which questions are answered the same way. (It does matter for belief, but that is not the current topic.) So I will assume that it is possible for Margit to rationally answer both questions the same way. (I will have much more to say about why this is a coherent assumption in Chapter 6.)

The way she should answer 6 is to take $A \wedge B$ as given. Hence she will take either option, Red-True or Blue-True, as being equivalent to just taking \$50, which she knows is the best she can do in the game. So in answering question 6, she will take it as given that both of these options are maximally good.

By hypothesis, she is answering question 5 and question 6 the same way. So she will take it to be part of the setup of question 5 that both options return a sure \$50. After all, that is part of the setup of question 6. But if she takes that as given, then conditionalising on either A or B does not change her expected returns. So now claims 2 and 3 are wrong; conditionalising on either conjunct won't make a difference because she treats each conjunct as given.

That is the totally general case. Assume that someone has competently deduced Y from X, and they know X. So they are entitled to answer the questions Q? and *If* X, Q? by the same method. Since the method for the latter takes X as given, so can the method for the former. So they can answer Q? taking X as given. What one can appropriately take as given

is closed under competent deduction. (Why? Because in the answer to Q? that starts with X, you can just go on to derive Y, and then see that it is also a way to answer *If* Y, Q?.) So they can answer Q? taking Y as given. So they can answer Q? in the same way they answer *If* Y, Q?.

So assuming there is no other reason to deny Single Premise Closure, adding a clause about how one may answer questions does not give us a new reason to deny it.

4.6.2 Multiple Premise Closure

That shows that IRT satisfies Single Premise Closure. The argument that it satisfies Multiple Premise Closure starts with the observation that Multiple Premise Closure more or less follows from Single Premise Closure plus a principle I'll call And-Introduction Closure.

And-Introduction Closure

If one knows some propositions, and one competently infers their conjunction from those propositions, while retaining one's knowledge of all those propositions, then one knows the conjunction.

Start with the standard assumption that a conclusion is entailed by some premises iff it is entailed by their conjunction. (It would take us way too far afield to investigate what happens if we dropped that assumption.) Given that assumption, in principle the only inferential rule one needs with multiple premises is And-Introduction. In practice, people do not generally reason via conjunctions in this way. Someone who knows $A \lor B$, and who knows $\neg A$, does not first infer $(A \lor B) \land \neg A$, and then infer *B* from that. They just infer *B*. It's a harmless enough idealisation, however, to model them as first inferring the conjunction whenever they use multiple premises. So I will assume that if I can show that IRT does not cause problems for And-Introduction Closure, and I've already argued that it does not cause problems for Single Premise Closure, then it does not cause problems for Multiple Premise Closure.

Here is the quick argument that IRT does not cause problems for And-Introduction Closure.

1. The key feature of IRT, the one that potentially causes problems for And-Introduction closure, is that one knows

that *p* only if one can take *p* for granted in one's current inquiry.

- 2. If, in the course of an inquiry, one knows some premises, then one can take them for granted in that inquiry.
- 3. If one can take some premises for granted in an inquiry, then one can take their conjunction for granted in that inquiry.
- 4. So, there is no IRT-based reason that And-Introduction Closure fails.

Premise 1 is just a restatement of my version of IRT, and premise 3 should be uncontroversial. If one can take some premises for granted, then one (rationally) is ruling out possibilities where they are false. To rule out possibilities where they are false just is to take their conjunction for granted. So those premises should be fairly uncontroversial. What is controversial is that the argument is sound, and, in particular, that premise 2 is correct.

The conclusion is not that Multiple Premise Closure holds. Maybe you think it fails for some independent reason, distinct from IRT. I don't think the other reasons that have been offered in the literature are compelling, but I am not building the failure of these reasons into IRT. So the main assumption behind the argument is that if adding the 'take for granted' clause to our theory of knowledge does not lead to closure violations, then nothing else in the theory does. The argument for that is basically that there isn't much more to the theory. So I think the argument is sound.

Still, it might look like the argument must be wrong. After all, it is easy to cook up cases where it looks like IRT leads to a closure failure. Here is one such example. It is another version of the Red-Blue game. I'll call the player in this version Morten. In this version, the red sentence is, once again, *Two plus two equals four*. This time the blue sentence is a conjunction *A and B*, where both *A* and *B* express historical facts that Morten has excellent, but not perfect, evidence for.¹³ Now the following four claims all seem true.

¹³ If you want to make this more concrete, pick a random history book off the shelf and choose two claims that are both reasonably specific—so there could easily be a mistake about the details—and not independently warranted.

- 1. Unconditionally, the only rational play is Red-True.
- 2. Conditional on *A*, the only rational play is Red-True. Even given *A*, playing Blue-True requires betting that *B* is true, and that's a pointless risk to run when playing Red-True only requires that two and two make four.
- 3. Conditional on *B*, the only rational play is Red-True. Even given *B*, playing Blue-True requires betting that *A* is true, and that's a pointless risk to run when playing Red-True only requires that two and two make four.
- 4. Conditional on $A \land B$, Blue-True is rationally permissible, and arguably rationally mandatory, since it weakly dominates Red-True.

So conditionalising on either one of *A* or *B* doesn't change anything, but conditionalising on $A \land B$ does change how Morten answers a question. So it looks like in this case Morten might know *A*, know *B*, and for all I've said be fully aware that these two things entail $A \land B$, but not know $A \land B$. So what's happened? How is this not a counterexample to premise 2?

The key thing to note is that when Morten is choosing what to do, the following things are all true about him.

- He can take *A* for granted. That is, he is rationally permitted to take *A* for granted in resolving his inquiry about what to do.
- Similarly, he can take *B* for granted.
- But he cannot both take *A* for granted and take *B* for granted. If both those things are taken for granted, then he can rationally infer that Blue-True will have a maximal payout, and hence that it is a rational play. And he cannot infer that.

It is cases like this one that required the clarification that I made at the end of Section 4.5. Morten here cannot take both of *A* and *B* for granted. So he doesn't know both those things. So this is not a case where he knows *A*, knows *B*, and doesn't know $A \wedge B$. Since he cannot take both *A* and *B* for granted, he does not know both of those things.

The picture I'm presenting here is similar to the picture Thomas Kroedel (2012) offers as a solution to the lottery paradox.¹⁴ He argues that we can solve the lottery paradox if we take justification to be a kind of permissibility, not a kind of obligation. And just as we can have individual permissions that don't combine into a collective permission, we can have individually justified beliefs that are such that we can't justifiably believe each of them. This isn't exactly how I'd put it. For one thing, I'm talking about knowledge not justification. For another, it's not that knowledge is a species of permission, as much as it behaves like permission in certain contexts, and those are just the contexts where counterexamples to And-Introduction Closure arise. These are minor points of difference though; I'm still basically relying on Kroedel's ideas.

Thinking of things the way Kroedel suggests helps say something positive about what is going on in this game. So far, I've said something negative—Morten does not know both that *A* and that *B*. That's enough to show that the case is not a counterexample to And-Introduction Closure. A counterexample would, after all, have to be a case where Morten knows both *A* and *B*. But saying what's not the case is not a helpful way to say what is the case. To say something more positive, it helps to think about other cases where permissions do not agglomerate. To that end, I'll talk through one case involving professional norms.

Professor Paresseux is, like most academics, in a situation where professional morality requires he do his fair share, but is fairly open about what tasks he does that will constitute doing his fair share. Right now he has two requests for work, R1 and R2, and while he is not obliged to do both, he is obliged to do at least one. So he may turn down R1, and he may turn down R2, but he may not turn down both. So as not to keep the reader in suspense, let's say up front that he is going to turn down

¹⁴ Different writers take different things to be *the* lottery paradox. In all cases, they concern what kind of non-probabilistic attitude an ideal agent would take towards the proposition that a particular ticket in a large, fair lottery will lose. It seems unintuitive to say that they will not believe this, since the ticket might win. And this will lead to an inconsistency, since they will believe of every ticket that it will not win, but also believe that a ticket will win. But if you say it is not belief, you seem to either get scepticism, or the view that the ideal agent can believe *p*, and not believe *q*, even though they think *q* is more probable than *p*. Which of the four problems I just mentioned is most salient to a writer tends to depend on their background commitments, but most people defend views on which at least one of the problems is genuinely problematic.

both. Our question will be, what exactly does Professor Paresseux do that's wrong?

To make this a little more concrete, and a little more complicated, I want to add two features to the case. First, accepting R1 would be better than accepting R2. He is uniquely well placed to do R1, and it would create more work for others if he turns it down. (As, indeed, he will.) But the norms governing Professor Paresseux are not maximising norms, and he does not violate them if he accepts R2 and rejects R1. Second, Professor Paresseux first turns down R1, let's say in the morning, and then later that day, let's say after a hearty lunch, turns down R2. Given that, there are three models we can have for the case, all of which have some plausibility.

The first model says that he was wrong to turn down R1. Here's a little argument for that, using language that seems natural. He should have accepted one of the requests. Since he was well placed to perform R1, it's also true that if he did one of them, it should have been R1. So he should have accepted R1, and turning it down was the mistake. Oddly, it turns out to have been made true that he did the wrong thing in turning down R1 by his latter decision to turn down R2, but that's just an odd feature of the case.

The second model says that this odd feature is intolerably odd. It says he was wrong to turn down R2. Here's a little argument for that. At lunchtime, he hadn't done anything wrong. True, he had turned down R1, but he had moral permission to do that. It was only after lunch that he made it the case that he violated a norm. So the violation must have been after lunch. So the violation was in turning down R2.

A third model says that both of these arguments are inconclusive. What's really true is simply that Professor Paresseux should not have turned down both requests. Which one individually was wrong? That, says the third model, is indeterminate. One of them must be, since he could not permissibly turn down both. But there is no fact of the matter about which it is.

If I had to choose, I would say that the third is the most plausible model. The arguments for the first two models are not terrible—indeed I think both are plausible models—but the arguments are equally compelling, and incompatible. So I suspect neither is entirely right. The third model, which says both of them are partially right—there is something not quite ok about both refusals—seems to better fit the scenario. But what I more strongly think is that each of these models is more plausible than either of the following two.

The fourth model is that there is a strong kind of agglomeration failure. It is determinately true that Professor Paresseux acted permissibly in turning down R1, and it is determinately true that he acted permissibly in turning down R2, but overall he acted impermissibly. It's true that in the abstract Professor Paresseux could have turned down each one. But in the particular context he is in—where these are the options to fulfil his duty to do his share of the work, and he does neither—is not one where he can (determinately) avail himself of both of these permissions.

The fifth model says that since he had to do his share and did not, and both refusals are ways of not doing his share, both of them are impermissible. This seems like overkill. It is much more intuitive that Professor Paresseux has done one wrong thing than that he has done two wrong things.

I hope I haven't traumatised too many readers with tales of people shirking professional responsibilities, because having Professor Paresseux's example on the table helps us lay out the options for what to say about Morten, given a further assumption about how he acts.

Morten plays the version of the Red-Blue game I just described, where the blue sentence is the conjunction of two plausible (and true) claims from a well-regarded history book he just read, and the red sentence is that two plus two is four. Let's assume that Morten looks at the rules, infers via his historical knowledge that playing Blue-True will have a maximal return, and so plays Blue-True. I think that this play is irrational, and if Morten knew the conjunction it would be rational, so Morten does not know the conjunction. But what do we say about Morten's knowledge of each conjunct? It turns out that there are five somewhat natural options that correspond to the five models I offered about Professor Paresseux. I'll simply list them here.

1. Morten knows the conjunct for which he has better evidence, and does not know the conjunct for which he has less good evidence. It was impermissible to take for granted the thing that was less well supported. This parallels the idea that Professor Paresseux did something wrong in turning down the request he was better placed to fulfil.

- 2. Morten knows the conjunct that he first took for granted, and not the conjunct that he took for granted second. When he first took one of the conjuncts for granted, that was a permissible mental act, but given that he had done it, it was impermissible to take the second for granted. This parallels the idea that whichever request Professor Paresseux turns down second is the impermissible turn-down, because it's then he becomes in violation of his duty.
- 3. It is indeterminate which conjunct Morten knows. He doesn't know both, because if he did then he could take both for granted, and he cannot take both for granted. Given both conjuncts, Blue-True is a rational play. So he must not know one, but there is no reason to say it is this one rather than that one, so it is indeterminate which he doesn't know. This parallels the indeterminacy solution to Professor Paresseux's puzzle.
- 4. Morten does know both conjuncts, since knowledge requires permissible taking for granted, and each of his takings for granted are individually permissible. But he doesn't know the conjunction, and so And-Introduction Closure fails.
- 5. Morten does not know either conjunct.

The fifth model seems like the least plausible. Somewhat unfortunately, it is also the model I defended (or at least committed myself to) in "Can We Do without Pragmatic Encroachment?". There I stated that knowledge requires conditionalising on what is known not to change any answers to interesting questions, and any question taken conditionally on an interesting proposition is itself interesting. So each of the questions *What should I play given the first conjunct is true*? and *What should I play given the first conjunct is true*? and *What should I play given the second conjunct is true*? are both interesting questions (in this technical sense of 'interesting'). Inquiring into the first question is incompatible with knowing the second conjunct. This was a fun way out the problem, but it was also overkill. Morten loses one bit of knowledge, not two, so my earlier view must be wrong.

Which of the other four models is correct? I think the fourth, which violates And-Introduction Closure, is the least plausible. That's largely

because it violates And-Introduction Closure. But the other three are all plausible, and are all consistent with And-Introduction Closure. (And note that all five are consistent with IRT. IRT itself says very little about this puzzle.) My preferred version of IRT says that typically the third option is correct—usually in cases like this it is indeterminate what is known.

There are mix-and-match options available. Perhaps if Morten's evidence for the first conjunct is (much) stronger than their evidence for the second conjunct, and it was the first one that they took for granted in reasoning, then they (determinately) know the first but not the second conjunct. I don't need to take a stance on whether cases like this ever arise to defend And-Introduction Closure. That's because all I need for any case like this is that one of the first three models be right. That can be true even if it is different models in different cases.

4.7 Summary

Putting all that together, IRT is consistent with Single Premise Closure and with And-Introduction Closure. Assuming that it is a harmless idealisation to treat anyone who uses multiple premises in reasoning as reasoning from the truth of all the conjunction of their premises, it follows that IRT is consistent with Multiple Premise Closure.

But this isn't quite the end of the story. Even if the arguments of the last two sections work, what they show is that there must be some way to explain away any apparent conflict between IRT and closure principles. The arguments do not, on their own, tell us what that explanation will look like, or whether it will have unacceptable consequences. Without such an explanation, we might be sceptical of the arguments of this chapter, and indeed of IRT itself. So I'll come back several times to issues about closure. In Chapter 6, I'll go over what IRT says about cases like Zweber's, and Anderson and Hawthorne's, more thoroughly.

Before I get to that though, it is time to say more about a notion that has done a lot of work so far but which has not been adequately investigated: inquiry.