

A/ Factorization

1. Internalism is, roughly, the idea that the internal and external can and should be factored apart in our theorizing.
2. Knowledge does not fit easily into an internalist picture; for knowledge is almost always partly internal and partly external.
3. Belief fits more easily into an internalist picture; for belief is (or anyway is often hoped to be) fully internal. One can believe that it is raining even if it is not raining.
4. Both knowledge and belief have mind-to-world direction of fit, unlike desire. And, according to orthodox internalist pictures:

... the general category for states with the mind-to-world direction of fit is belief. The belief is true if it fits the world, false otherwise. Although true and false belief are the same mental state in different worlds, the place of belief in the economy of mental life depends on its potential connection with truth. Knowledge is merely a peculiar kind of true belief. Since Gettier showed that even justified true belief is insufficient for knowledge, epistemologists have expended vast efforts attempting to state exactly what kind of true belief knowledge is, but the programme is assigned no significance for the philosophy of mind. On such a view, knowledge is to be explained in terms of belief, and belief is what matters for the understanding of mind. The converse attempt to explain belief in terms of knowledge sounds eccentric and perverse. To summarize the orthodoxy: belief is conceptually prior to knowledge. (p. 2)

The idea that belief is conceptually prior to knowledge has another source: the internalist conception of mind, and world external to mind, as two independent variables. Belief is simply a function of the mind variable. Truth is simply a function of the external world variable, at least when the given proposition is about the external world. For the internalist, knowledge is a function of the two variables, not of either alone; whether one knows that it is raining does not depend solely on one's mental state, a state which is the same for those who perceive the rain and those who hallucinate it, but it also does not depend solely on the weather, a state which is the same for those who believe the appearances and those who doubt them. The internalist therefore conceives knowledge as a complex hybrid crying out for analysis into its internal and external components, of which belief and truth respectively are the most salient. (p. 5)

5. One sees internalism in attempts to define knowledge. Knowing that it is raining is the conjunction of an internal state—something like: justifiedly believing that it is raining—and an external state—something like: being where it is raining.
6. Attempts to define knowledge have not been successful.
7. Knowledge has entailments. Perhaps knowledge entails true belief. But one cannot always find a remainder.

Although being coloured is a necessary but insufficient condition for being red, we cannot state a necessary and sufficient condition for being red by conjoining being coloured with other properties specified without reference to red. Neither the equation 'Red = coloured + X' nor the equation 'Knowledge = true belief + X' need have a non-circular solution. (p. 3)

B/ Knowledge-First

8. A different approach starts with knowledge.
9. Instead of thinking of knowledge in terms of belief, we think of belief in terms of knowledge.

... to believe that p is to treat p as if one knew p —that is, to treat p in ways similar to the ways in which subjects treat propositions which they know. (pp. 46-7)

10. *Analogy*: A fish flops on the deck. What is it doing? It is trying to swim away, something it also does in the good case in which it swims away. But what is it to try to swim away? One could think that we should understand trying to swim away in terms of swimming away—trying to swim away is acting in a way that is relevantly similar to the way that something that is swimming away is acting.
 - a. What the fish is doing contributes to its flourishing in the good case, and, on this view, we make sense of what it is doing both in the good and the bad case by appeal to the connection between what it is doing and the flourishing it helps manifest in the good case.
11. *Similarly*: A plastic apple is on the table. A hungry someone reaches toward it. Why is she doing that? She believes that the thing on the table is an apple. But what is it to believe that the thing on the table is an apple? One could think that we should understand believing in terms of knowing—believing that there is an apple on the table is treating the proposition *that there is an apple on the table* as if one knew it, e.g. being disposed to reach across the table upon being hungry.
12. Here is a very simple theory of belief that has this sort of structure:

- a. Internal state s of some object carries the information that p just if, necessarily, if conditions are epistemically normal, and the object is in state s , then p . (Like knowing, carrying the information that p is factive.)
 - b. To believe that p is to have some internal state s that carries the information that p .
 - c. Hence: to believe that p is to be in a state that in the good case carries the information that p .
13. Something similar could be done with justification. More on that later in the semester.

C/ Is There Enough There, There, Internally?

- 14. Suppose that the internal supervenes on one's intrinsic physical state.
- 15. It may be that rational behavior supervenes on the internal—it may be that there could no change in how it is rational for one to behave without some change to their intrinsic physical state. Still, we can explain one's rational behavior by citing contentful attitudes, which do not supervene on their intrinsic physical state.

When we attribute mental states to each other in ordinary language, the conditions of which we speak are often broad. That one sees Naples, that one remembers Naples, that one keeps referring to Naples—all are broad conditions, because none obtain in cases in which one lacks even indirect causal connection with Naples, whereas one's internal physical state has no such necessary dependence on a city. Similarly, that one loves Mary and that one hates Mary are broad conditions, for they depend on a relation to the particular individual named. (pp. 52-3)

16. Why is one writing that note? Because one loves Mary.

When I attribute a mental state to you in ordinary language, the implications of my statement can easily outrun your boundaries. I say that you see paper; as every sceptic knows, you could be in the same internal state as someone who sees paper without seeing paper yourself: my statement is true only if paper is there before your eyes, outside you. In some sense, my statement is not purely about you. For theoretical purposes, would it not be more perspicuous to resolve the mixture into its underlying elements, by separating a statement purely about you from another purely about the environment external to you? After all, causation is local—no action at a distance—so does not the causal explanation of your actions require the isolation of what is local to you from background conditions on the environment? This resolution might amount to an analysis, giving necessary and sufficient conditions for the truth of my original statement.

Alternatively, it might replace that statement without being equivalent to it, by doing its causal-explanatory work better. Either way, internal and external factors are separated. Such a picture is internalist. (p. 47)

17. But is there a level of contentful explanation of rational behavior that is internalist? Not obviously.

The internalist is obliged to concede that content ascriptions in natural language express broad rather than narrow conditions, but nevertheless insists that they consequently fail to reflect the structure of the underlying facts. On this view, such ascriptions characterize the subject by reference to a mixture of genuinely mental states and conditions of the external environment. The challenge to such an internalist is to make good this claim by isolating a level of description of contentful attitudes that is both narrow and genuinely mental, not merely neurophysiological. If there is such a mixture of the internal and the external, it should be possible to separate out its constituents. The broadness of content ascription in natural languages shows that the required level of description does not simply lie to hand, but must be constructed; its effect is therefore to put the burden of proof on the internalist. (pp. 53-4)

18. A parallel issue arises about 'externalist' attitudes, such as knowledge:

Parallel considerations apply to internalism about the attitudes themselves. Factive propositional attitudes are a source of blatantly broad conditions, whether or not their contents are broad. Even when the sentence 'One believes that A' does not express a broad condition, the conditions expressed by 'One knows that A', 'One sees that A', and 'One remembers that A' are almost always broad. While conceding this, the internalist nevertheless insists that such constructions fail to reflect the structure of the underlying facts. Factive constructions are held to characterize the subject by reference to a mixture of genuinely mental states and conditions on the external environment. As before, the challenge to the internalist is to make good this claim by isolating a level of description that is both narrow and genuinely mental. The effect of the broad natural language semantics is against to put the burden of proof on the internalist. (pp. 54)

D/ Explanation

19. Knowledge appears to play a role in the causal explanation of action that belief does not.

Consider a causal explanation as simple as 'He dug up the treasure because he knew that it was buried under the tree and he wanted to get rich'. Note that the explanandum ('He dug up the treasure') makes reference to objects in the environment (the treasure) as well as to the subject's immediate physical movements. The internalist cannot substitute 'believe' for 'know' in the explanation without loss, for the revised explanans, unlike the original, does not entail that the treasure was where he believed it to be; the connection between explanans and explanandum is therefore weakened. (p. 62)

20. Knowledge appears to play a role in the causal explanation of action that true belief does not.

A burglar spends all night ransacking a house, risking discovery by staying so long. We ask what features of the situation when he entered the house led to that result. A reasonable answer is that he knew that there was a diamond in the house. To say just that he believed truly that there was a diamond in the house would be to give a worse explanation, one whose explanans and explanandum are less closely connected. For one possibility consistent with the new explanans is that the burglar entered the house with the true belief that there was a diamond in it derived from false premises. For example, his only reason for believing that there was a diamond in the house might have been that someone told him that there was a diamond under the bed, when in fact the only diamond was in a drawer. He would then very likely have given up his true belief that there was a diamond in the house on discovering the falsity of his belief that there was a diamond under the bed, and abandoned the search. In contrast, if he *knew* that there was a diamond in the house, his knowledge was not essentially based on a false premise. Given suitable background conditions, the probability of his ransacking the house all night, conditional on his having entered it believing truly but not knowing that there was a diamond in it, will be lower than the probability of his ransacking it all night, conditional on his having having entered it knowing that there was a diamond in it. In this case, the substitution of 'believe truly' for 'know' weakened the explanation, by lowering the probability of the explanandum conditional on the explanans. The substitution of 'believe' without 'truly' for 'know' would do even worse. (p. 62)

21. Suppose that rational behavior supervenes on beliefs and desires and, to control for as many confounds as possible, suppose that belief and desire somehow are individuated internally. Then one would not have to cite what the burglar knows in order to explain

why he rationally spent all night ransacking the house. Still, citing that the burglar knew that there was a diamond in the house might improve the explanation.

22. Perhaps the belief-based explanation will cite many things the burglar does and does not believe. (E.g., the burglar spent all night ransacking the house because the burglar believed ... and did not believe ...). But generality is a good-making feature of explanation.

Good explanations have an appropriate generality. If one cites a sufficient condition for the condition to be explained, or one near enough so for the purpose in hand, the purported explanation can nevertheless fail because the condition to be explained would still have obtained in the same way even if the cited condition had not obtained. For example, one can explain why someone died by saying that he was run over by a bus; the explanation becomes worse, not better, if one specifies that the bus was red, for its colour had nothing to do with the death. (p. 81)

23. Perhaps the belief-based explanation will cite a long disjunction, each disjunct being many things the burglar might believe and not believe. (E.g. the burglar spent all night ransacking the house because (the burglar believed ... and did not believe ...) or (the burglar believed ... and did not believe ...) or But relevance is a good-making feature of an explanation.

... if someone was crying because she was bereaved, it does not improve the explanation to say that she was crying because she was bereaved or chopping onions. (p. 83)

24. Something that I will return to later in the semester: sometimes we explain rational behavior, not by citing a cause, but by citing a basis. Compare: jumping because someone said jump and startled you and jumping because someone said jump and you are following orders.

E/ Primeness

25. Seeing water is prime. Cases α and β may be cases of seeing water. Case γ may be internally like α and externally like β , yet not be a case of seeing water.

What is the point of classifying a case in which one sees water separately from a case in which one is in exactly the same internal physical state but sees only a mirage? The difference may not matter for one's action at the next instant, if the action is itself individuated by its internal physical nature. If it were individuated

broadly, we should still be wondering about the point of broad individuation. But our interest is not confined to action at the next instant—if there is one; if time is dense, there is not. One is thirsty; how likely is one to be drinking soon? Likely enough, if one sees water. Much less likely, if what one sees is a mirage. Even if drinking is individuated narrowly, its explanation in terms of the earlier state of the system involves the presence of water in the environment, not just the earlier internal physical state of the agent. Concepts of broad mental conditions give us a better understanding of connections between present states and actions in the non-immediate future, because the connections involve interactions with the environment. (p. 75)

26. Knowing that it is raining is prime. Cases α and β may be cases of knowing that it is raining. Case γ may be internally like α and externally like β , yet not be a case of knowing that it is raining.

When the causal explanation depends on the primeness of (D_1 or E_1) and (D_2 or E_2) and ..., as section 3.4 argued that it often will, the extractable narrow condition D_1 or D_2 or ... typically plays no explanatory role; it is a sort of epiphenomenon. What would give the narrow condition an explanatory role is compositeness, not analysability; the argument for primeness therefore tells against such an explanatory role for the narrow condition. (p. 92)