

24.500 spring 05  
topics in philosophy of mind  
session 6

- teatime

self-knowledge

# plan

- Shoemaker, matters arising
- transmission again
- Peacocke & Co.

## argument 2 (§V)

- A. self-blind speaker George will recognize the paradoxical character of 'P but I don't believe that P'
- B. since he is rational, this recognition will lead him to avoid Moore-paradoxical sentences
- C. further, George will recognize that he should give the same answer to 'do you believe that P?' and 'P?'
- D. there is nothing in his behavior, verbal or otherwise, that would give away the fact that he lacks self-acquaintance
- E. if George really is self-blind, then perhaps we are too!
- F. it seems better to take this as a reduction ad absurdum of the view that self-blindness is a possibility

- another way of stating the conclusion of the argument:

if

- rational intelligent George has the conception of the various mental states and can entertain the thought that he has this or that belief, wants to find out about his mental life, etc.,

then [modulo a later qualification]

- after a period of experimentation\*, George will have self-knowledge that is similar in extent to ours, and that he can attain in the sorts of circumstances in which we can attain self-knowledge

\*George is of normal intelligence, etc., not a genius.

- B** since he is rational, this recognition will lead him to avoid Moore-paradoxical sentences
- why, since he can't follow rules like 'if you have such-and-such intentions that make appropriate an assertive utterance of 'p', don't assert 'I don't believe that p''?
  - he can (avoid M-p sentences), since he has the assertive use of language
  - why, though?

- what I need to support my case is a reason for thinking that someone with first-order beliefs plus [normal conceptual capacity and rationality] would thereby have the use of language

- George will be able to understand language
- he could see that certain ends he has would be furthered by saying certain things, so he will say those things
  - Grice-1 intention: audience should believe p (via this intention)
  - Grice-2 intention: audience should believe the speaker believes p
- the argument so far shows that George will have (something close to) the practice of assertion, with Grice-1 intentions

- OK, but what about a case where there would be a point in speaking with Grice-2 intentions? (when the audience already believes that p)
  - George would have a motive for saying 'p':
    - it's true that p, so:
    - I should act as if I believed that p
    - in this instance so acting would mean saying 'p', because plainly this would have good consequences for me

how can this be right? cf. 'p, so Smith should act as if he believed that p'



- a suggestion:
  1. the audience needs to know whether I believe that p
  2. now, it's true that p
  3. so, if I assert that p, the audience will come to know that I believe that p
  4. as a result of the above reasoning (given that I want to help the audience), I will assert that p
- but the step from (1) and (2) to (3) is suspect—I can reach (3) only if I know (or have good reason to believe) that I believe that p
- and this can't be assumed, because I am supposed to be self-blind

- what about ‘tell me some things you believe’? wouldn’t that leave George speechless?
- no—he could treat it as equivalent to the request ‘tell me some things that are true’
- so far, step D is supposedly looking good; now, a new pair of objections

- D there is nothing in his behavior, verbal or otherwise, that would give away the fact that he lacks self-acquaintance
1. George will be puzzled, because ‘p and I don’t believe that p’ might be warranted by the evidence
    - this case is not really conceivable [is this right?]
  2. George will not seem to have self-acquaintance with his past beliefs
    - this objection does require me to qualify my claims [can’t the objection be met? explain “But if...” at the top of p. 44]

1. George will be puzzled, because ‘p and I don’t believe that p’ might be warranted by the evidence
  - “the total third-person evidence concerning what someone believes about the weather should include what evidence he has about the weather—and if it includes the fact that his total evidence about the weather points unambiguously toward the conclusion that it is raining, then it cannot point unambiguously toward the conclusion that he doesn’t believe that it is raining”

- “the total third-person evidence concerning what someone believes about the weather should include what evidence he has about the weather”
- my total evidence: the cat is wet, the weather forecast is for rain, I am going out without my umbrella, carrying important papers that will spoil if it’s raining...[so, I don’t believe it’s raining]
  - this is evidence that I have, it does not include the fact that this is my evidence (for that includes facts about my beliefs—that I believe that the cat is wet, etc.)
- “and if it includes the fact that his total evidence about the weather points unambiguously toward the conclusion that it is raining...”
- but it doesn’t

- a suggestion: my total evidence includes the fact that I would say that it's raining, if asked
  - my total evidence: the cat is wet, the weather forecast is for rain, I am going out without my umbrella, carrying important papers that will spoil if it's raining...
  - and: I would say that it's raining, if asked
    - so, my total evidence does not support the conclusion that I don't believe that it's raining
- but: am I really entitled to "if I were asked, I would say that it's raining" without self-knowledge?
  - I believe that it's raining; I want to help the questioner; so, if I were asked, I would say that it's raining

## extending argument 2

- Moore-analogues:
  - let us do  $X$ !—but I don't want to do  $X$  (cohortative)
  - would that  $P$  be true!—but I don't hope that  $P$  is true (optative)
  - I'll do  $A$ , but I don't intend to
    - not straightforward to extend the previous argument to cover these cases
    - e.g. “tell me some of your desires”
    - we will return to this later

## a few points

- the account leaves out experiences, sensations, occurrent thoughts, etc.
- privileged access is not explained (just peculiar access)
- the “as-if he has”-to-“he has” move is suspicious
- OK, someone could have self-knowledge by reasoning in George’s way; what about us?
- an insight (to be taken up later): self-knowledge is economical



# transmission revisited: a general problem

- 1 I am thinking that water is wet
- 2 if I am thinking that water is wet, then water exists

hence:

C water exists

- “We have already seen that Sally has empirical knowledge that her environment contains water” (Brown, 235); see also Brewer and Sawyer, quoted in Davis, Externalism and armchair knowledge (but see Davis, Externalism...)
- but: Sally could know (1), (2), and yet not be in a position to know (C) empirically (imagine she has grown up in a community that thinks ‘water’ is like ‘phlogiston’; if we like, we can add that against all the evidence, she believes that water exists, is found in the oceans, etc.); if so, the strategy cannot preserve closure
  - this is even more evident in the “architecturalism” case

# conscious attitudes, attention, and self-knowledge

- topic: conscious thoughts, conjectures, decisions, etc.
- claim: there are certain “constitutive features” of these “states” [events?], “which can be identified by broadly philosophical investigation”

# part I

- consciously recalling p is (always?) “subjectively different” from consciously recalling (or seeming to consciously recall) a distinct proposition q
- activities like trying to remember something can “occupy your attention” [?]
- these activities are not “objects of attention”, because they are not “experienced”
- the kinds of attention involved in conscious thought and perception are subspecies of “a single, general kind of attention”

# attention

- the psychological use of 'attention' is a semi-technical term
- capacity limits: amount of information processing resources available for a task
- selection: keeping capacity fixed, the selection of a particular object/region/property etc. on which processing is concentrated (see 70)

BLUE		GREEN
RED		BLUE
GREEN		RED
BLUE		BLUE
YELLOW W		RED
GREEN		YELLOW W

- if what I have said about the occupation of attention by conscious thought is along the right lines, then the occupation of attention at least in directed thought also performs a function of selection
  - is this more than the observation that what one thinks about is to some extent under voluntary control?
  - and is it the mental episode on which one's attention is directed, or its topic? (see Martin, 103, 119)
- without such selection, human thought would be chaotic
  - how does that follow?

## part II

1. an apparent memory that Peacocke taught at NYU
  2. a judgment that Peacocke taught at NYU
  3. a self-ascription of the belief that Peacocke taught at NYU
- the “mental event” described in (2) might be one’s “reason” for making the self-ascription described in (3)

[I]n making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks me “Do you think there is going to be a third world war?,” I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question “Will there be a third world war?”. (Evans 1982, 225)

- Peacocke's point: the memory example is a “special case of Evans' procedure, rather than any kind of rival to it”
  - Martin seems to interpret Peacocke differently at 114

# “failure of pseudo-transitivity”



1. an apparent memory that Peacocke taught at NYU
  2. a judgment that Peacocke taught at NYU
  3. a self-ascription of the belief that Peacocke taught at NYU
- (1) is not a reason for (3), although one may achieve self-knowledge by moving straight from (1) to (3)
  - mystifying



## part III

- against:
  - the “no-reasons” account (derived from Shoemaker)
  - internalist introspection
  - the spurious trilemma (recall Boghossian: either observation, inference, or nothing)

# Boghossian's paradox

1. we know our own thoughts immediately (i.e. not by inference)
2. either non-inferential self-knowledge is based on some form of inner observation or on nothing—at any rate, on nothing empirical
3. since you cannot tell by mere inspection of an object that it has a given *relational* or *extrinsic* property, inner observation is not the source of self-knowledge
4. self-knowledge is based on nothing, and so is not a cognitive achievement (from 1, 2, 3)

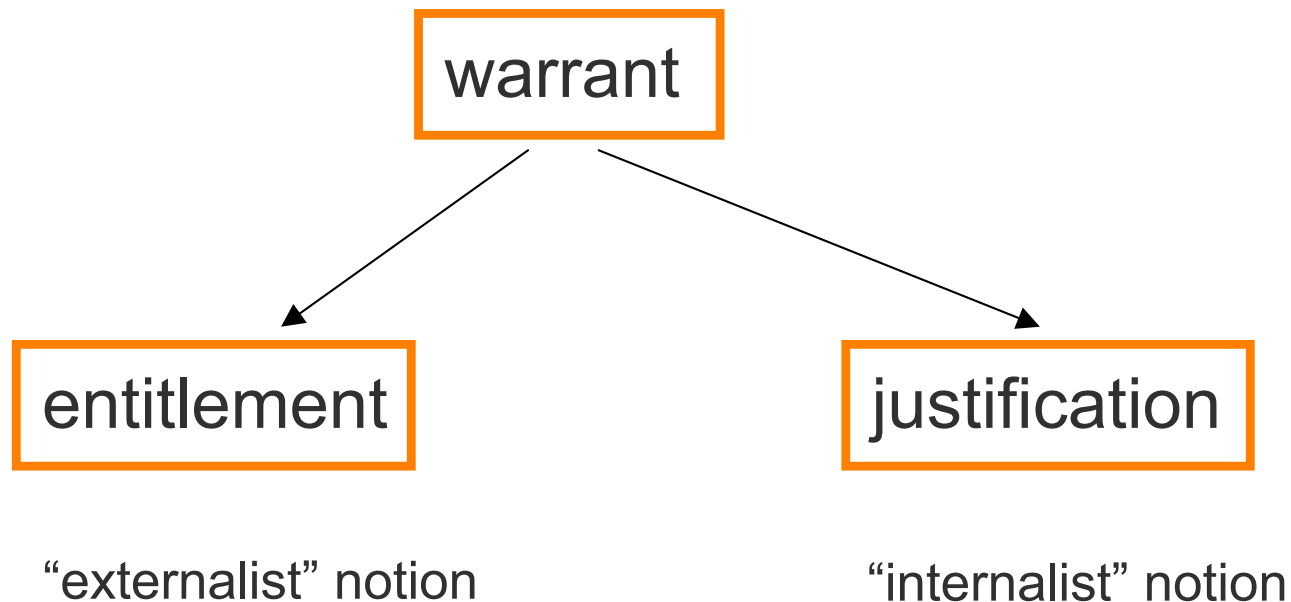
but this conclusion is absurd

- Peacocke: (2) is false, because one's experience can "make reasonable" the second-order judgment that one is having an experience
- contrast our earlier response, namely that the conclusion is not absurd

- no reasons: never a personal-level, causal, reason-giving explanation of why a thinker believes he believes that p
- natural no reasons epistemology: reliabilism
  - is that because it doesn't give "personal level" explanations?—and what are these?
  - note the quote from Shoemaker on 77 does not express his own view
- the no-reasons view is false because conscious states "can give reasons" (83)
  - note that they "are reasons" earlier
  - it is somewhat unclear what Peacocke's talk of "reasons" amounts to
  - see Unger, *Ignorance*, for some considerations in favor of taking reasons to be known propositions
    - TO BE CONTINUED...

# Our entitlement to self-knowledge

- what epistemic warrant do we have to make judgments about our mental states?
  - Burge's answer: "entitlement"



- where does the entitlement derive from?
- not from reliability

rather, from:

- the role of the relevant judgments in critical reasoning, and:
- a constitutive relation between the judgments and their subject-matter—or between the judgments about one's thoughts and the judgments' being true

the line of thought, briefly drawn:

1. to be capable of critical reasoning...some mental states and acts must be knowledgeably reviewable
2. the entitlement [associated with this knowledgeable reviewability] must be stronger than that involved in perceptual judgments
3. there must be a non-contingent, rational relation, of the sort to be explained, between relevant first-person judgments and their subject matter or truth

- critical reasoning is reasoning that involves an ability to recognize and effectively employ reasonable criticism or support for reasons and reasoning
- as a critical reasoner, one recognizes reasons as reasons
- clearly, this requires a second-order ability to think about thought contents or propositions...
  - so far, not stipulated that a critical reasoner has to be able to think about her mental states

- animals and small children reason non-critically, as do we, much of the time
- to be a fully critical reasoner, one must be able to identify, distinguish...propositions as denied, hypothesized
- so critical reasoning requires thinking about one's thoughts
- but it further requires that that thinking be normally knowledgeable
- the interest lies in the ground of [this] requirement; why must we be normally knowledgeable about our thoughts when we reflect upon them?

apparently, we have established:

1. to be capable of critical reasoning...some mental states and acts must be knowledgeably reviewable
  - but this seems to be just a trivial consequence of the way the notion of critical reasoning was introduced



- why must we be normally knowledgeable about our thoughts when we reflect upon them?
- I will answer this question in three stages
  1. to evaluate reasoning critically, one must have an epistemic entitlement to one's judgments about one's thoughts
  2. critical reasoning requires that one know one's thoughts...
  3. this knowledge must take a distinctive, nonobservational form

regarding stage 2, it seems we haven't established:

1. to be capable of critical reasoning...some mental states and acts must be knowledgeably reviewable



1. to evaluate reasoning critically, one must have an epistemic entitlement to one's judgments about one's thoughts
  - since one's belief about one's thoughts is an integral part of critical reasoning, one must have an epistemic right to it
  - to be reasonable in the whole enterprise one must be reasonable in that essential aspect of it
  - ...so one must have an epistemic entitlement to one's judgments about one's attitudes

- OK, but why is this warrant entitlement?
- why isn't it justification?

- 2 critical reasoning requires that one know one's thoughts...
- for, if not, reflection [beliefs about one's attitudes, etc.] would not add a rational element to the reasonability of the whole process of critical reasoning

- why, given that one might reasonably believe p without knowing p?

- a brute error is an error that indicates no rational failure and no malfunction in the mistaken individuals, e.g. belief caused by a perceptual illusion
- simple observational model: one's epistemic warrant always rests partly on a pattern of veridical[?], but brute, contingent, non-rational{?} relations—which are plausibly always causal relations—between one's attitudes and judgments about one's attitudes

\*presumably: the attitude causes the appropriate judgment

3 this knowledge must take a distinctive, nonobservational form

- if it didn't, "brute errors" would be possible [in every case?]
- but they aren't possible [in every case?]

- surely "brute errors" are possible for self-ascriptions of seeing that p
- and if so, we haven't explained the "distinctive nonobservational" way we know we see that p
- according to the Rylean, brute errors in self-ascription are possible (just as brute errors in other-ascription are possible)
- has anything been said that shows why Ryleanism is false (or is incompatible with the presence of critical reasoning)?
- and—setting Ryleanism to one side—why should we agree Burge's claim about brute error?

- next time:
  - Peacocke and Burge contd., selections from Wittgenstein, *Blue Book*; Shoemaker, Self-reference and self-awareness; selections from Evans, *The Varieties of Reference*