

T E M I

AUTOCONOSCENZA

di Patrizia Pedrini

ABSTRACT – Come conosciamo la nostra mente? Che tipo di rapporto abbiamo con i nostri stati mentali? Siamo gli spettatori degli eventi mentali che popolano la nostra mente oppure abbiamo con essi un rapporto in prima persona? L'opposizione al Cartesianesimo e l'affermazione del naturalismo in filosofia della mente hanno coinciso con il trionfo di una prospettiva in terza persona, ovvero di carattere osservativo, verso i nostri stati mentali. Vi sono invece ragioni fenomenologiche evidenti per sostenere che il rapporto che abbiamo con i nostri stati mentali non è principalmente di tipo spettatoriale e osservativo.

1. INTRODUZIONE
2. IL PROBLEMA DELL'AUTOCONOSCENZA
3. IL MODELLO CARTESIANO E IL MODELLO PERCETTIVO DELL'AUTOCONOSCENZA
4. IL MODELLO PRATICO DELL'AUTOCONOSCENZA
5. BIBLIOGRAFIA

1. INTRODUZIONE

Come conosciamo la nostra mente? Che tipo di rapporto abbiamo con i nostri stati mentali?

Siamo gli spettatori degli stati mentali che popolano la nostra mente oppure abbiamo con

essi un rapporto immediato, “in prima persona”? L’opposizione al Cartesianesimo e l’affermazione del naturalismo in filosofia della mente hanno coinciso con il trionfo di una prospettiva “in terza persona”, cosiddetta, verso gli stati mentali di cui siamo portatori: secondo questa prospettiva, quando conosciamo i nostri stati mentali, non facciamo altro che osservarli, nello stesso modo in cui osserviamo gli stati mentali di un altro o gli eventi del mondo. A questo esito hanno contribuito anche l’influenza delle riflessioni critiche di Ludwig Wittgenstein e di Gilbert Ryle, l’eredità del movimento psicoanalitico e i nuovi risultati della psicologia sperimentale e delle scienze cognitive. Inoltre, siamo esseri capaci di conoscere i nostri stati mentali autorevolmente? Detto altrimenti, quando possediamo uno stato mentale di prim’ordine (cioè uno stato mentale costituito da un atteggiamento verso una certa proposizione p), siamo in grado di formare credenze di second’ordine vere e giustificate (cioè atteggiamenti di credenza che abbiamo per oggetto gli stati mentali di prim’ordine), oppure, come già voleva Jean-Paul Sartre, e come la scienza cognitiva e la vita ordinaria in generale sembrano testimoniare, spesso succede che gli altri siano invece i migliori giudici di noi stessi, condannati come siamo all’ignoranza di larghe porzioni della nostra vita mentale o a forme di autoinganno?

Una delle posizioni maggiormente accreditate in letteratura sostiene che vi sono importanti ragioni fenomenologiche per affermare che il rapporto che abbiamo con i nostri stati mentali, in condizioni normali, non è principalmente di tipo spettatoriale e osservativo, bensì è in larga parte un rapporto che può essere definito “in prima persona”, ovvero raggiunto in modo prevalentemente immediato e non inferenziale, a differenza di quello che succede per la conoscenza degli stati mentali altrui.¹ Secondo questa posizione, fenomeni

¹ Ho già sostenuto una parte di queste tesi in Pedrini [2009], a cui mi permetto di rimandare.

come l'autoinganno o l'ignoranza circa i propri stati mentali non pregiudicano l'esistenza di fenomeni paralleli di piena autorità del soggetto che si autoconosce.²

Per comprendere in che modo una parte centrale della conoscenza che abbiamo dei nostri stati mentali si distingue dalla conoscenza che abbiamo degli stati mentali altrui, e per definire con chiarezza le sue caratteristiche epistemiche e metafisiche della prima, occorre ripercorrere l'itinerario concettuale offerto negli ultimi decenni dal dibattito sull'autoconoscenza, fino agli esiti più recenti della ricerca.

2. IL PROBLEMA DELL'AUTOCONOSCENZA

Il problema filosofico centrale del dibattito sull'autoconoscenza è fornito da un fenomeno universale e manifesto, eppure di non facile spiegazione teoretica. Il fenomeno in questione è il seguente: sembra che siamo in grado, almeno per larga parte della nostra vita mentale cosciente, di conoscere in maniera *autorevole* e *immediata* i nostri stati mentali. L'autorità con cui riferiamo quali sono molti degli stati mentali che intratteniamo e l'immediatezza con cui li conosciamo, come vedremo, segnano quell'asimmetria di base tra la condizione del soggetto, che conosce molta parte della sua vita mentale "in prima persona" (questa l'espressione tecnica ormai entrata nell'uso), e quella dell'interprete, il quale apprende invece gli stati mentali di un altro "in terza persona".

Che significa che conosciamo un gran numero dei nostri stati mentali in modo autorevole e immediato? In generale, diciamo che una conoscenza è *autorevole* quando colui che la consegue la ottiene in virtù della speciale *posizione epistemica* nella quale si trova rispetto al contenuto di questa conoscenza. In un'ampia varietà di casi ordinari, il portatore di uno

² Ho sviluppato questa linea di ricerca in Pedrini, "On the Failure of Authority", manoscritto.

stato mentale mostra di conoscerlo con un' autorità di fronte alla quale un interprete di solito si rimette. Questo sembra suggerire che il portatore di uno stato mentale conosce “meglio” di un interprete questo stato, che ne ha, come si dice, un “*accesso privilegiato*”³ - privilegiato in virtù, appunto, della sua posizione epistemica. Per esempio, nessuno come me stesso può sapere cosa penso in questo momento, quali credenze ho rispetto alla situazione politica italiana, quali intenzioni nutro circa le mie azioni di oggi pomeriggio, o cosa di preciso desidero che avvenga entro i prossimi due mesi. Anche supponendo che un interprete possa in qualche caso indovinare quali desideri a mia insaputa possiedo, o a quali mire profonde sono rivolte certe mie attività, quale sia il loro sfondo motivazionale ecc., ciò non toglie che in qualsiasi momento io possa fare in modo che egli non sappia o dubiti - mentre io continuo a sapere e non dubito - che cosa ho in mente esattamente. *Autorità* significa, perciò, l'unicità e superiorità, in termini di precisione, accuratezza e affidabilità, dello status epistemico delle mie opinioni su me stesso e sui contenuti della mia vita mentale, rispetto alle opinioni che un interprete può formarsi su questi stessi contenuti.

In secondo luogo, diciamo che una conoscenza è *immediata* quando, appunto, non è conseguita in maniera “mediata”, cioè quando non è ottenuta per il tramite di inferenze condotte a partire dall'evidenza disponibile.⁴ Nel caso degli stati mentali, normalmente essi offrono evidenza di sé attraverso il comportamento, verbale e pratico, del loro portatore. Un interprete conosce gli stati mentali dell'interpretando inferendoli dall'evidenza in questione (quel che l'interpretando dice, fa ecc.); li conosce cioè “mediatamente”. L'interpretando, invece, ne ha solitamente una conoscenza che non si avvale di alcuna evidenza e non segue

³ Per la nozione di accesso privilegiato, si vedano Alston [1971], Heil [1988] e la raccolta di articoli a cura di Wright, Smith & Macdonald [1998]. Per gli esordi del dibattito sul tema dell'autorità in prima persona, si veda anche Davidson [1984; 1987].

⁴ Si veda ad esempio Moran [2001]

il metodo dell'interpretazione. *Immediatezza* qui significa, perciò, che la conoscenza della maggior parte dei nostri stati mentali ordinari (credenze, intenzioni, desideri ecc.) non è il risultato di una “ricostruzione *autointerpretativa*”,⁵ cioè di una strategia di attribuzione a noi stessi di stati mentali la cui esistenza sarebbe indipendente dalla conoscenza che ne abbiamo, nonché individuabili a prescindere dall'identificazione di essi come *nostri*. Faccio un esempio: prendiamo una qualsiasi credenza ordinaria che *p*, che razionalmente ho formato e intrattengo, come la credenza che la chiave di casa si trova nella mia tasca destra. Per ascrivermela, cioè per sapere che credo che la chiave si trova nella mia tasca destra, non procedo ad attribuirgliela a me stesso in seguito al riconoscimento di un corpo di evidenze sufficienti a stabilire che c'è una persona che crede che la chiave si trovi nella sua tasca destra e che questa persona è identica a me stesso. Questo è il lavoro dell'interprete, ordinario o clinico. Se mia madre vuol sapere se credo che la chiave si trovi nella mia tasca destra, deve chiedermelo, oppure osservarmi mentre la cerco (come pure riconoscermi). In questo modo soltanto, mia madre è in grado di inferire che credo che la chiave si trova nella mia tasca destra. Allo stesso modo, anche un analista, pur incaricato di rivelare all'analizzando parti di sé che non conosce, non potrebbe neppure dare inizio a un'analisi, se non disponesse delle parole del paziente, accanto ai suoi gesti: solo l'analizzando può cominciare a dire sotto quale descrizione ha compiuto tali gesti, come si sente e cosa le circostanze gli suscitano, esprimendo tra l'altro in questo modo la conoscenza concettuale che ne possiede. Su questa base, l'analista procede ad attribuirgli una serie di stati mentali, naturalmente non senza aggiustamenti critici progressivi. Nel caso dell'autoconoscenza in prima persona, invece, l'autoascrizione e, quindi, la conoscenza di un individuo *S* della sua credenza

⁵ La definizione è di Taylor [1977].

che la chiave si trova nella sua tasca destra, quando effettivamente ha questa credenza, o che si sente tradito da suo fratello, quando così si sente, è conseguita da *S* immediatamente e automaticamente.

Il fenomeno dell'autoconoscenza dei nostri stati mentali appena descritto, per quanto universale e manifesto, è tuttavia terribilmente problematico per la riflessione filosofica. Le ragioni principali di difficoltà a darne una spiegazione sono due. In primo luogo, su una spiegazione coerente di esso pesa quella che potremmo chiamare la “*minaccia cartesiana*”: sostenere che *S* ha conoscenza immediata e automatica della credenza, intenzione o desiderio che *p*, che intrattiene, è una tesi che potrebbe apparire in larga parte simile, se non addirittura identica, alle tesi cartesiane di trasparenza e infallibilità della mente, secondo le quali ogni volta che *S* esemplifica una certa proprietà mentale, non può non saperlo.⁶ Questa equazione, e la sua duratura influenza, sono di fatto responsabili di molta parte dello scetticismo contemporaneo verso l'autoconoscenza, come capacità distinta dalla (e in generale più affidabile della) conoscenza in terza persona. Il fatto è che, rispetto al tempo in cui Descartes concepì la sua dottrina, disponiamo di dati scientifici (e di rimbalzo, come effetto della nuova educazione scientifica, forniti anche dalla letteratura e dalla vita di ogni giorno), che dimostrano che nessuno di noi è pienamente trasparente a se stesso e infallibile.⁷ È divenuto luogo comune che possediamo stati mentali di cui siamo inconsapevoli, ignoranti,

⁶ Si veda Heil [1988], p. 238, che si riferisce a questa tesi come “tesi di incorreggibilità” della conoscenza dei nostri stati mentali.

⁷ Questa idea è stata resa famosa dagli esperimenti di Nisbett e Wilson, in Nisbett and Wilson [1977]. Si veda anche Nisbett e Ross [1980]. Nel set sperimentale descritto nell'articolo di Nisbett e Wilson, ad esempio, al soggetto vengono attribuite credenze di cui è ignorante, ma che emergono da altri dati comportamentali non verbali. Un problema spinoso che potrebbe essere sollevato contro gli esperimenti di Nisbett e Wilson è che l'attribuzione al soggetto sperimentale di queste credenze non è corretta, almeno secondo certi criteri per l'attribuzione di credenze. I dati comportamentali segnalano la presenza di un qualche stato mentale di cui il soggetto non è consapevole e che emerge a livello di comportamento, ma non è ovvio che si tratti di credenze propriamente dette. Lo spazio a disposizione non consente una discussione approfondita di questa obiezione.

sulla cui natura e contenuto possiamo essere in errore. Così, assieme alle nuove rappresentazioni della mente, le nostre intuizioni sono cambiate, e sembra che di questi stati inconsci si possa al massimo avere la conoscenza fallibile o incompleta che un interprete potrebbe avere⁸. Niente, allora, in principio, sembra stabilire la speciale autorità epistemica di un particolare individuo sulla sua mente. Se esistono infatti stati mentali di cui siamo ignoranti, un interprete potrebbe invece conoscerli. Se è un esperto, potrebbe accedervi più facilmente di noi, o accedervi *tout court*. La sua autorità epistemica sembra dunque superiore alla nostra; e la via ricostruttiva, in terza persona, più affidabile. Ciò confligge direttamente con la tesi dell'autorità della prima persona che ho esposto sopra nelle sue linee essenziali.

Ora, avanzare la tesi di immediatezza, invece, *non* è un altro modo per affermare che per ogni stato mentale che *S* possiede, data la natura della mente di *S*, *S* sa di intrattenerlo. Essa non è una tesi metafisica sulla natura dell'io, bensì una tesi *epistemologica* su un modo della consapevolezza, ancorché si tratti di un modo niente affatto contingente, bensì centrale e vitale per la salute psichica dell'individuo. La delimitazione non è triviale. Ciò su cui di fatto si regge l'obiezione contenuta in quella che ho chiamato “minaccia cartesiana” è l'imputazione, in funzione anticartesiana, di una generalizzazione della tesi di immediatezza a qualsivoglia stato mentale. L'obiettore ci mostra che nella classe degli stati mentali stanno plausibilmente anche stati che sono veri di noi e di cui non siamo consapevoli, o la cui conoscenza potrebbe essere solo ricostruttiva. In questo modo, la tesi generale di immediatezza è diventata invalida per sovrabbondanza di controesempi. L'obiezione manca il punto. Essendo l'immediatezza una tesi su un modo della consapevolezza, se uno stato mi

⁸ La nozione di stato mentale “inconscio”, notoriamente, ha un importantissimo pedigree filosofico e psicologico, che non è possibile riassumere nel corso di questo articolo. Mi limito a segnalare che, nonostante sia una nozione cardinale nella teoria psicanalitica, nella filosofia della mente contemporanea vi sono autori, come ad esempio Searle [1992], che sono piuttosto scettici circa l'esistenza di tali stati.

è ignoto, per definizione non lo conosco, né *a fortiori* lo conosco immediatamente. Ora, anche con i difetti di completezza e accuratezza che ciascuno di noi registra nella propria conoscenza di sé, è manifesto che le credenze, intenzioni, desideri che attivamente formiamo in ogni istante ci sono noti di norma in modo immediato, e non attribuzionalmente né inferenzialmente. La tesi di immediatezza come tesi epistemologica su un modo della consapevolezza non intende affatto escludere che vi siano stati di cui non abbiamo conoscenza. È una teoria che si incarica di illustrare in che *modo* conosco *quegli* stati che conosco e, in particolare, che cosa significa conoscerli dal punto di vista della prima persona, piuttosto che da quello della terza. Non pretende, insomma, di essere una teoria che spieghi perché non conosco gli stati inconsci, né d'altra parte li nega.⁹

Vi sono, tuttavia, alcune osservazioni che meritano di essere fatte sulle ragioni per le quali certi stati mi sono ignoti. Le cause che mi impediscono di conoscere un certo stato mentale immediatamente non sono chiare in ogni singolo caso, ma vi sono certamente alcune circostanze generali che favoriscono o generano ignoranza. È intelligibile, per esempio, che la tesi di immediatezza possa non applicarsi a fatti che riguardano verità psicologiche strutturali, morali o qualitative in genere, della nostra persona (come ad esempio la disposizione al senso di colpa, oppure la tendenza ad essere affettivamente dipendenti dagli altri ecc.) . Per il coglimento di esse potremmo mancare delle risorse sia concettuali sia motivazionali (o le une o le altre) necessarie alla formazione del relativo atteggiamento proposizionale. Non è infatti necessario che io abbia riflettuto sulle mie azioni o comportamenti in maniera che potremmo definire auto-osservativa. Né, qualora lo faccia, tale riflessione è sufficiente a che acquisti conoscenza, perché potrei appunto non riuscire ad e-

⁹ Si veda Moran [2001], capitolo1, e Pedrini [2009], Introduzione.

sprimerla concettualmente. Stati complessi, oppure insiemi di atteggiamenti con storie causali intricate, nonché disposizioni come tratti del carattere o orientamenti psicologici generali della persona, come anche stati di tipo subpersonale (cioè uno stato che non sia coglibile attraverso le categoria della psicologia di senso comune, detta anche di livello personale, come ad esempio uno stato neurale), possono quindi ben sfuggire, in ragione della finitezza epistemica umana, alla nostra conoscenza immediata, se non addirittura essere per loro natura inconoscibili (è il caso degli stati subpersonali).¹⁰ E, di fatto, il modo immediato di autoconoscenza non è invariabilmente utilizzato dalle persone per conoscere ogni verità psicologica che le riguarda. Non si deve del resto pensare che la tecnica di conoscenza in terza persona sia appannaggio esclusivo di una terza persona. Si tratta, anzi, di una tecnica in molti casi plausibilmente impiegabile da parte di ciascuno di noi per il conseguimento di una parte interessante di conoscenza di noi stessi, che non sarebbe altrimenti ottenibile. Questa parte di autoconoscenza può infatti divenire oggetto di opinione da parte del soggetto che la esemplifica solo per via evidenziale e ricostruttiva. Quella di immediatezza, perciò, è una tesi assai più modesta delle tesi cartesiane di infallibilità e trasparenza. Essa afferma non più dell'enunciato minimo che quando conosco, per esempio, la mia credenza che la chiave di casa si trova nella mia tasca destra, cioè quando so di possedere questa credenza, indipendentemente dal fatto che il suo contenuto sia vero (cioè che sia vero che la chiave di casa effettivamente si trova nella mia tasca destra), lo so senza che questa conoscenza sia in alcun modo una *conquista cognitiva*, quale sarebbe per esempio la scoperta *a posteriori* dell'identità empirica "acqua=H₂O", oppure la scoperta che un'altra mente, quella di mia madre, poniamo, crede che la chiave di casa sia nella mia tasca destra.

¹⁰ Per una più ampia discussione su questi punti, si veda Pedrini [2009], capitolo 2.

Vi è poi una seconda difficoltà che pesa sulla spiegazione dell'immediatezza, che potremmo denominare “*minaccia teoretico-epistemologica*”. Essa consiste nella seguente circostanza: anche con le delimitazioni sopra fornite, la tesi di immediatezza rimane un'idea provocatoria, che necessita di ulteriore analisi: sembra infatti naturale domandarsi come può l'autoconoscenza qualificarsi come conoscenza genuina, non arbitraria, e anzi *sostantiva*, pur non essendo basata su evidenza. Il problema che la tesi di immediatezza qui solleva è quale sia la *fonte* dell'autorità con cui esprimo il contenuto della mia mente, la *base* a partire dalla quale conseguo tale autorità – una base che potrei citare, in seguito, come giustificazione della *conoscenza* che vanto. Il fatto è che, se qualcuno mi chiede, in spirito scettico, come faccio a sapere che davvero credo che la chiave sia nella mia tasca, non avrei una risposta soddisfacente. Potrei al massimo ribadire che lo so, il che non accontenterebbe certo un interlocutore scettico, che mi chiedesse quali evidenze sono in grado di fornire per giustificare questo sapere. Ora, anche non nutrendo un particolare senso del dovere verso gli standard fissati da uno scettico, resta difficile spiegare come la mia conoscenza dei miei stati mentali possa non essere basata su evidenza. Se non è basata su evidenza, sembra in effetti che debba essere basata “su niente”;¹¹ e che, dunque, non se ne possa offrire una giustificazione.

3. IL MODELLO CARTESIANO E IL MODELLO PERCETTIVO DELL'AUTOCONOSCENZA

Come ho già accennato più sopra, l'autoconoscenza è stata spiegata da Descartes in una famosa dottrina sulla natura della mente riflessiva, nota come “dottrina del cogito”.¹² Essa

¹¹ L'espressione è di Boghossian, in Boghossian [1989], p. 485.

¹² Si confronti Descartes [1641], Meditazione II e VI.

comprende almeno due tesi fondamentali, che qui ci interessano direttamente, e che vale la pena di definire con chiarezza prima di procedere:

- (1) la tesi di trasparenza della mente a se stessa;
- (2) la tesi di infallibilità del giudizio sui propri stati mentali;

La tesi (1) afferma che per ogni proprietà mentale M esemplificata da un individuo S , S non può non sapere di esemplificare M . La tesi (2) è una conseguenza di (1), e afferma che per ogni giudizio che S si forma sulle proprietà mentali M che esemplifica, questo giudizio non può essere falso.

E' divenuto impossibile sottoscrivere acriticamente entrambe le tesi, come si è visto. Non possiamo non tenere nel conto i macroscopici controesempi all'idea di mente trasparente e infallibile che la vita quotidiana, la psicologia, le scienze cognitive ci offrono. Le persone si sottopongono ad analisi per scoprire cosa provano, hanno stati mentali che emergono nei test sperimentali, ma di cui non sono consapevoli (si veda nota 7), e così via. Dopo tre secoli di dominio cartesiano, il tema dell'“accesso privilegiato” del soggetto alla propria mente è diventato un problema filosofico, che ha meritato di essere trattato in maniera sistematica, prevalentemente perché nuove correnti di pensiero e nuove scuole psicologiche ci hanno fornito ragioni per pensare che la mente sia *assai* diversa da come Descartes l'aveva descritta. Sembra che si debba abbandonare l'idea di trasparenza, perché siamo ignoranti su un gran numero di stati mentali che ci riguardano, e per la tesi di infallibilità le cose non vanno meglio, data l'esistenza sia di fenomeni di ignoranza dei nostri stati menta-

li, sia di fenomeni di autoattribuzione di stati mentali che di fatto non abbiamo. Il modello cartesiano di autoconoscenza deve dunque considerarsi in bancarotta.

Ci viene a questo punto suggerita una via di uscita: nulla di completamente originale, risale ad un'epoca persino precedente alla filosofia di Descartes, e ha tra i suoi iniziatori Hume [1739-1740] e Locke [1690];¹³ e tuttavia, è una via d'uscita non senza attrattive, poiché apparentemente in grado di appianare sia il problema dell'immediatezza sia quello dell'autorità. Quando espose la dottrina del cogito, Descartes non si limitò, in effetti, solo ad asserire le tesi di trasparenza e infallibilità. Disse di più, ne fornì una spiegazione, riprendendo una teoria della coscienza diffusa tra gli empiristi della sua epoca.¹⁴ Per sostenere, in particolare, la tesi di infallibilità, egli dette una caratterizzazione del tipo di consapevolezza che abbiamo dei nostri stati mentali come di una "percezione interna", imputabile a un senso interno. Ancor più specificamente, disse che questa percezione interna, che avremmo dei nostri stati mentali, è una relazione con il contenuto della nostra mente che viene conseguita per il tramite di un "organo", tipicamente percettivo, appunto, che indicò con la metafora di un "occhio interno", capace di "vedere" gli avvenimenti che si muovono nel "teatro" della nostra mente. La nozione di "introspezione" entra in gioco precisamente qui: come l'etimo della parola suggerisce, gli esseri umani - o meglio, la mente, la "sostanza pensante", che a rigore potrebbe non essere la stessa cosa che un essere umano - avrebbero la capacità di ispezionare, cioè di guardare e dirigere la propria attenzione percettiva verso gli accadimenti che nella loro mente occorrono. (Si noti che non mi sto concentrando su nessuna teoria specifica della percezione. Mi interessa invece avvalermi di uno stereotipo di percezione sufficientemente generale, ma anche sufficientemente rigoroso, da

¹³ Per un inquadramento bibliografico del problema, si vedano Di Francesco [1998] e Nannini [2002].

¹⁴ Si veda Crapulli [1988].

comprendere un certo numero di caratteristiche che qualsiasi teoria della percezione presuppone).

Le caratteristiche che la percezione interna dei propri stati mentali avrebbe, qualora il modello percettivo dell'autoconoscenza fosse il modello corretto da seguire, sono state elencate, in chiave critica, da Sidney Shoemaker nella prima e seconda di una serie di tre famose lezioni [1994a, 1994b, 1994c], nelle quali ha difeso con vigore la tesi che è un grave errore filosofico quello di concepire l'autoconoscenza come conoscenza di tipo percettivo: la relazione che ciascuno di noi ha con i propri stati mentali non è una relazione del tipo di quella che intratteniamo con fatti percepibili del mondo esterno, come il fatto che fuori piove, o che c'è un albero di fronte a noi, o che nostra madre ha i capelli biondi. Non è l'immagine corretta di autoconoscenza, perché al nostro stereotipo di percezione occorrono le seguenti otto condizioni, che l'autoconoscenza non soddisfa:

- i) che si possieda un *organo* della percezione (interna);
- ii) che si abbiano *esperienze* percettive;
- iii) che nell'esperienza percettiva sia rappresentato un *oggetto* percepibile, da cui eventualmente derivare un giudizio *fattuale* nel quale l'oggetto sia coinvolto;
- iv) che si disponga di informazioni per l'*identificazione* dell'oggetto percepito;
- v) che l'oggetto abbia *proprietà intrinseche*, cioè non relazionali;
- vi) che l'*attenzione* sia coinvolta, o guidi, o spieghi, l'atto della percezione;
- vii) che la percezione sia *causalmente* prodotta dall'oggetto;
- viii) che l'oggetto che ne è causa esista *indipendentemente* dal conoscente.

Vediamo che succederebbe se questo fosse il modello adatto per descrivere la nostra relazione con i contenuti nella nostra mente; vediamo, cioè, come dovremmo descrivere la nostra mente affinché un modello di autoconoscenza che soddisfacesse queste condizioni fosse vero. Prendiamo come esempio un qualsiasi atteggiamento proposizionale. Cominciamo da quello classico, la credenza. Innanzitutto, lo statuto ontologico di un organo della percezione mentale, come sarebbe l'occhio (o senso) interno cartesiano, è a dir poco dubbio, se non, al limite, persino inintelligibile: non solo non lo rinveniamo, né con la riflessione né con l'indagine empirica sui nostri organi di senso, ma è decisamente oscuro il modo in cui dovrebbe funzionare, quali "percezioni" dovrebbe fornire, su quali oggetti dovrebbe appuntarsi.¹⁵ Facciamo un esempio. Prendiamo la credenza che Giorgio Napolitano è il presidente della repubblica italiana. Stiamo supponendo che essa sia l'oggetto sul quale l'*attenzione* del nostro *organo* di senso interno dovrebbe dirigersi per percepirlo; la sua esistenza sarebbe *indipendente* dal conoscente, cioè da noi stessi; l'oggetto sarebbe inoltre *allocato* in qualche *luogo* della nostra mente, e rimarrebbe *sconosciuto* finché non fosse colto percettivamente; data la posizione epistemica favorevole dal percipiente (nella quale già si trova o che è stata raggiunta con un qualche esercizio della volontà), la percezione di esso dovrebbe dar adito, per il tramite di un impatto *causale* sull'organo della percezione interna, a un'*esperienza* percettiva di un qualche tipo, durante la quale il percipiente *identificherebbe* questo *oggetto* che sarebbe identico alla sua stessa credenza che Giorgio Napolitano è l'attuale presidente della repubblica, cogliendone le *caratteristiche intrinseche* di

¹⁵ Il caso sarebbe diverso se prendessimo in considerazione la propriocezione. Nella propriocezione, l'oggetto dei nostri stati di dolore o piacere sarebbero le parti del nostro corpo interessate. Armstrong [1963] tende ad estendere il caso della propriocezione a tutti quanti gli stati mentali, compresi gli atteggiamenti proposizionali. Quello che ho argomentato in Pedrini [2009] è invece una visione degli atteggiamenti proposizionali tale da escludere che abbia senso una sovrapposizione di essi al caso della propriocezione. Non è possibile nello spazio disponibile discutere questo tema più estesamente.

tale credenza. L'inesistenza di un organo introspettivo potrebbe non ancora risultare fatale all'implausibilità del modello – potremmo candidare il cervello, oppure le forme di propriocezione, come fa Armstrong [1963]. Mentre invece sono fatali al modello le altre condizioni, e in particolare (vii)-(viii), che la nostra conoscenza che crediamo che Giorgio Napolitano è l'attuale presidente della repubblica (assieme a numerosissime altre) non soddisfa. Partiamo proprio da (viii): affinché sia soddisfatta la condizione di indipendenza dell'oggetto della percezione dal percipiente (d'ora in poi, seguendo Armstrong [1963; 1968], “tesi di indipendenza”[*distinct existence argument*]), dovremmo raffigurarci la credenza che Napolitano è il presidente come un fatto della nostra mente (o addirittura un “oggetto” non fattuale) per sua natura conoscibile *a posteriori*, sotto forma di “scoperta”. Non sapremmo, cioè, fino al momento in cui la nostra attenzione epistemica non fosse adeguatamente diretta verso questo oggetto o fatto, che crediamo che Napolitano è il presidente. Si deve pertanto supporre che questa credenza si sarebbe formata indipendentemente da noi, indipendentemente, cioè, dalla nostra coscienza di essa, e sarebbe sopravvissuta alla nostra ignoranza. Tra l'altro, a rigore, non si spiegherebbe neppure come la nostra attenzione potrebbe volontariamente dirigersi verso essa, se è sconosciuta. Dovremmo figurarci un caso fortunato, di serendipità. Questa ulteriore conseguenza controintuitiva dipende dal fatto che gli oggetti, o fatti, indipendenti che incontriamo non solo hanno un impatto causale su di noi (o sui nostri organi di senso), ma anche, normalmente, generano una credenza percettiva, il cui contenuto rappresenta l'oggetto coinvolto *in quanto* identificato. Non escludo che vi siano esperienze percettive dotate di contenuto non proposizionale, il cui impatto sulla nostra esperienza dia adito a stati “subdoxastici”,¹⁶ e certamente, però, quando

¹⁶ Un stato subdoxastico dovrebbe essere uno stato la cui funzionalità assomiglia al consueto stato di creden-

così fosse, per definizione, nessun oggetto verrebbe propriamente identificato – cioè, non si formerebbe alcuna credenza percettiva del tipo che vedo che c'è un albero di fronte a me. Ora, che significa che *identifico* la credenza che Napolitano è il presidente della repubblica? Dovrebbe significare che, tra la miriade di oggetti mentali che popolano la mia mente, ne distinguo uno in particolare, *dimostrativamente*, le cui caratteristiche *intrinseche*, non relazionali, mi permettono di identificarlo come la credenza che Napolitano è il presidente della repubblica. Ma non esiste nulla che assomigli alle caratteristiche intrinseche, non relazionali, della credenza che Napolitano è il presidente della repubblica. Per caratteristica intrinseca intendiamo una caratteristica posseduta dall'oggetto indipendentemente da qualsiasi relazione con l'ambiente o altri fatti e oggetti, e che sia coglibile da parte di un percipiente, dotato di adeguati concetti, dopo ispezione dell'oggetto percepito. Se qualcosa del genere fosse concepibile, dovremmo tra l'altro confrontarci con una grave evenienza: la possibilità di errori d'identificazione. Volevamo certamente spiegare l'immediatezza senza cadere nell'ingiustificabilità, e congiuntamente spiegare la fallibilità in funzione anti-cartesiana; ci siamo, così, imbarcati nel test di un modello che prometteva di spiegare, contemporaneamente, immediatezza, ignoranza ed errore. Potrebbe a questo punto sembrare che l'obiettivo sia stato perfettamente colto: gli stati mentali sono stati indipendenti dal conoscente, se possiamo sbagliarci su di essi, o se possiamo esserne ignoranti, per poi magari, ad un certo punto, scoprirli. Eppure, la situazione eccezionalmente problematica in cui il modello percettivo ci mette riguarda proprio il genere di errori che sulla base di questo modello sarebbe possibile commettere e prevedere, e il tipo di lacune che potremmo

za, ma che manca di coscienza esplicita e di un contenuto proposizionale noto al suo portatore. La discussione sulla natura e il ruolo degli stati subdoxastici occupa intensamente la scienza cognitiva. Per una panoramica generale della collocazione di questi stati all'interno della teoria della mente si veda Marconi [2001] e Marraffa [2003].

immaginare di avere. Poiché un atteggiamento proposizionale è composto dal contenuto dell'atteggiamento e dal tipo di atteggiamento, potrei, per esempio, commettere un errore di identificazione della stato mentale *relativamente al contenuto*:

- potrei, cioè, erroneamente scambiare la credenza “*che Napolitano è presidente della repubblica italiana*” con la credenza “*che Vladimir Putin è il presidente della repubblica russa*”.

Oppure, potrei commettere un errore di identificazione dello stato mentale *relativamente all'atteggiamento*:

- potrei, cioè, erroneamente scambiare la credenza “*che Napolitano è presidente della repubblica italiana*” con il *desiderio* o l'*intenzione* “*che Napolitano è presidente della repubblica italiana*”.

Entrambe queste eventualità sono controintuitive. Il Modello Percettivo non è la strada giusta. Ma qual è, allora, l'alternativa percorribile?

4. IL MODELLO PRATICO DELL'AUTOCONOSCENZA

Il modello che nello spazio disponibile potrò solo cominciare a delineare tende a considerare la connessione tra certi stati mentali (non tutti, si noti bene!) e la conoscenza che il soggetto ne possiede come una *connessione a priori*, di natura concettuale. Non è per un caso che siamo “reporter” affidabili delle nostre rabbie, paure, credenze, intenzioni, e il ri-

levamento di questa connessione *a priori* è parte della critica all'assimilazione dell'introspezione a una percezione interna. Perché la connessione è necessaria? La sua natura *a priori* si spiega sulla base della *razionalità* che è specifica di soggetti umani, dotati di psicologia riflessiva: quando credo che Napolitano è il presidente della repubblica italiana riesco a conoscere questa credenza immediatamente perché *io stesso* l'ho formata e la mantengo. E l'ho formata e la mantengo non già sulla base del fatto che l'ho in qualche modo individuata nella mia mente, bensì sulla base delle *ragioni* che ho per credere che Napolitano è il presidente della repubblica. L'autoconoscenza immediata e autorevole di questi stati mentali non è quindi una capacità "in più", che ci troviamo ad avere, di consapevolezza circa fatti psicologici che ci riguardano e che, venisse a mancare o non fosse per qualche motivo esercitata, lascerebbe inalterati i fatti psicologici in questione. La disanalogia con i giudizi che formiamo sul mondo esterno, dove i fatti conosciuti hanno condizioni di verità logicamente indipendenti dall'effettuazione del giudizio su di essi, incoraggia a imboccare la strada di un modello completamente diverso da quello puramente teoretico ed epistemologico che applichiamo al mondo esterno: si tratta di un modello noto nella filosofia anglo-americana come "Modello Pratico-Deliberativo". In esso, l'attenzione teorica è rivolta alla persona come agente epistemico, invece che alla persona come presunto spettatore di fatti indipendenti. Le *capacità epistemiche dell'agente* si esercitano sulle ragioni che egli ha per formare certe credenze, e non sul cogliimento di credenze che dovremmo concepire come preesistenti e logicamente indipendenti da lui.

Il carattere speciale dell'autoconoscenza in prima persona non può quindi essere più inteso come "accesso epistemico" a un regno speciale al quale solo una persona, me stesso, ha diritto di ingresso. Viene invece pensato, in questo modello, in termini di *responsabilità*

verso la propria mente che il soggetto ha in virtù del fatto che quella mente è la *sua*. Gli atteggiamenti di quest'ultimo, le sue credenze ecc., sono viste da lui come espressive delle relazioni razionali, ovvero comprensibili, che intercorrono tra lui e il suo ambiente, e non come la mera successione di rappresentazioni incomprensibili di cui, per qualche ragione, è l'unico testimone. Si tratta, quindi, di consapevolezza immediata, e non osservativa, tale da implicare il riferimento a sé tramite ciò che è sotteso all'uso del pronome "io", e non per il tramite di una descrizione sotto cui la persona potrebbe anche non riuscire ad identificarsi.¹⁷

Richard Moran [2001] ha presentato questo modello in maniera estesa. Alla "conoscenza attraverso la scoperta" (*knowledge by discovery*) Moran oppone un'altra forma di conoscenza, più adatta all'autoconoscenza: la "conoscenza attraverso decisione" (*knowledge by resolution*, Moran [2001], cap. 4). La decisione (*resolution*) di cui parla Moran altro non è che la valutazione delle ragioni a favore o a sfavore di un atteggiamento. Valutando le ragioni a favore o sfavore di un atteggiamento e dando il mio assenso a quelle che ritengo conclusive, non farei altro, secondo Moran, che esercitare una sensibilità razionale, di tipo propriamente *pratico*, su cosa *dovrei* credere, data l'evidenza che ho a disposizione. Quest'idea è il fondamento, nell'opera di Moran, del Modello Agenziale dell'autoconoscenza. Egli propone, perciò, che si accantoni l'immagine dell'autoconoscete come passivo annotatore di stati mentali formatisi indipendentemente da lui stesso e dall'esercizio delle sue capacità epistemiche di decisore razionale rispetto all'evidenza, e che si passi a un'immagine di autoconoscenza nella quale, per spiegare immediatezza e autorità, si focalizzi proprio sulla sensibilità razionale dell'agente epistemi-

¹⁷ Si vedano ancora Perry [1977; 1979], Shoemaker [1963; 1968] e Campbell [1995].

co. Così si spiega anche l'autorità della prima persona: non solo credo che Napolitano è il presidente della repubblica, ma necessariamente so anche di crederlo, poiché io stesso ho formato tale credenza in virtù delle ragioni che ho per formarla.¹⁸ Moran pensa anche che una "domanda deliberativa" "Devo credere/intendere/desiderare, ecc. che p?" sia sempre logicamente applicabile ad un atteggiamento proposizionale, perché questi atteggiamenti sono il risultato di una deliberazione sull'evidenza disponibile.

Se questa teoria può spiegare con una certa facilità immediatezza e autorità, non è però esente da problemi. Ecco una sintetica panoramica delle principali obiezioni che è possibile muovere ai dispositivi teorici proposti da Moran:

- Se il modello pratico-deliberativo fosse corretto, rischieremmo di stare propugnando una visione gravemente iper-intellettualizzata dell'autoconoscenza, e della mente più in generale. Sembra infatti che la maggior parte dei nostri stati mentali sia frutto di una deliberazione e che sia questa loro caratteristica ciò che li rende conoscibili da parte del soggetto che ne è portatore. Tuttavia, è ovvio che le sensazioni, ad esempio, non sono prodotte da alcuna deliberazione, e tuttavia le conosciamo generalmente in maniera altrettanto immediata e autorevole. Per far fronte a questa obiezione, i fautori del modello pratico-deliberativo, come Moran, cercano di rivendicare la legittimità della restrizione del campo di applicazione della loro teoria ai soli atteggiamenti deliberati, dimostrando che non è ne-

¹⁸ Burge [1996] e Peacocke [1996], discutendo il problema dell'*entitlement*, ovvero della fonte dell'autorità epistemica verso la conoscenza dei nostri pensieri, hanno espresso concezioni simili.

cessaria una teoria unificata dell'autoconoscenza. In altre parole, per le sensazioni dovrebbe essere elaborata un'altra teoria, di tipo non pratico-deliberativo.¹⁹

- D'altra parte, sembra che vi siano anche altri stati proposizionali la cui componente deliberativa non è così ovvia. Ad esempio, se desidero mangiare un gelato, non sembra che abbia conseguito questo atteggiamento proposizionale sulla base di alcuna valutazione, deliberativa o meno, di alcuna evidenza. Alla luce del genere di difesa che i propugnatori del modello pratico-deliberativo elaborano, questa obiezione sembra più difficilmente affrontabile. Occorrerebbe qui una terza teoria, da accostare a quella che spiega l'autoconoscenza delle sensazioni e al modello pratico-deliberativo. Se anche un modello unitario non sempre è una virtù teoretica degna di essere conseguita, la proliferazione di teorie appare come una manovra *ad hoc*.

- Infine, sembra plausibile sostenere che abbiamo molte credenze che non sappiamo di avere, cioè che sia possibile credere che p , senza contemporaneamente credere che crediamo che p , mentre il modello in questione assicura che vi è una connessione *a priori* tra la credenza che p e la sua autoconoscenza, che invece in questi casi sembra mancare.²⁰ Questo fenomeno sembra richiedere che si debba restringere la teoria dell'autoconoscenza solo ai casi di conoscenza esplicita dei nostri stati mentali. Non è chiaro, allo stato attuale della ricerca, che natura potrebbe avere una conoscenza non esplicita dei propri stati mentali.

¹⁹ Si veda perlomeno Finkelstein [2003]. Per altre obiezioni a Moran, si veda anche O'Brian [2003] e Bagnoli [2007].

²⁰ Si rinvia a Pedrini [2009], capitolo 4, per un maggior approfondimento di questi temi. Altre varianti di modelli promettenti, che invocano l'agente epistemico, ma senza porre l'accento sulla manovra deliberativa, si trovano per esempio in Bilgrami [2006] e Coliva [2009; 2001]. Si veda anche, tra le risorse enciclopediche on-line, la voce "Self-Knowledge" in *Stanford Encyclopaedia of Philosophy*.

BIBLIOGRAFIA

Alston, W. P. (1971), “Varieties of Privileged Access”, *American Philosophical Quarterly*, 8, pp. 223-241.

Armstrong, D. M. (1963), “Is Introspective Knowledge Incorrigible?”, *Philosophical Review*, 72, 4, pp. 417-432.

Bagnoli, C. (2007), “The Authority of Reflection”, *Theoria*, 58, pp. 43-52.

Bilgrami, A. (2006), *Self-Knowledge and Resentment*, Harvard University Press, Cambridge, Massachusetts, London, England.

Bernecker, S. and Dretske, F. (2000), *Knowledge: Readings in Contemporary Epistemology*, Oxford University Press, Oxford.

Boghossian, P. (1989), “Content and Self-Knowledge”, in *Philosophical Topics*, 17, pp. 5-26; rist. in Bernecker and Dretske (2000), pp. 479-498.

Campbell, J. (1995), *Past, Space, and Self*, MIT Press, Cambridge, Ma.

Coliva, A. (2009), “Self-Knowledge and Commitments”, *Synthese* 171, pp. 365–375.

Coliva, A. (2011) "One Variety of Self-Knowledge: Constitutivism as Constructivism", in A. Coliva (ed.) *The Self and Self-Knowledge*, Oxford University Press, Oxford, forthcoming.

Crapulli, G. (1988), *Introduzione a Descartes*, Laterza, Roma-Bari.

Davidson, D. (1984), “First Person Authority”, in *Dialectica*, 38, pp. 101–112; reprinted in Davidson, D. (2001), *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford, 2nd edn (trad.it. *Verità e Interpretazione*, Il Mulino, Bologna)

Davidson, D. (1987), “Knowing One’s Own Mind”, in *Proceedings and Addresses of the American Philosophical Association*, 61, 441–58; reprinted in Davidson, D. (2001), *Su-*

jective, Intersubjective, Objective, Clarendon Press, Oxford (trad. it. *Soggettivo, intersoggettivo e oggettivo*, Raffaello Cortina, Milano)

Descartes, R. (1641), *Meditazioni metafisiche*, trad. it. in *Discorso sul metodo e meditazioni metafisiche, con le obiezioni e risposte*, Roma-Bari, Laterza, 1978, voll. 1 e 2 (e *Opere*, a cura di G. Cantelli, Milano, Mondadori, 1986, pp. 197-266).

Di Francesco, M. (1998), *L'io e i suoi sé. Identità personale e scienza della mente*, Raffaello Cortina, Milano.

Finkelstein, D. H. (2003), *Expression and the Inner*, Harvard University Press, Cambridge, Ma.

Heil, J. (1988), "Privileged Access", *Mind*, New Series, 97, 386, pp. 238-251.

Hume, D. (1739-1740), *Trattato sulla natura umana*, a cura di E. Lecaldano, in *Opere filosofiche*, vol. 1, Laterza. Roma-Bari, 1987.

Locke, J. (1690), *Saggio sull'intelligenza umana*, trad. it. C. Pellizzi (1972), Laterza, Roma-Bari.

Marconi, D. (2001), *Filosofia e scienza cognitiva*, Laterza, Roma-Bari.

Marrassa, M. (2003), *Filosofia della psicologia*, Laterza, Roma-Bari.

Moran, R. (2001), *Authority and Estrangement. An Essay on Self-Knowledge*, Princeton University Press, Princeton, NJ.

Nannini, S. (2002), *L'anima e il corpo. Un'introduzione storica alla filosofia della mente*, Laterza, Roma-Bari.

Nisbett, R. and Ross, T. D. (1980), *Human Inference: Strategies and Shortcomings of Social Judgement*, NJ, Prentice Hall.

Nisbett, R. and Wilson, T. (1977), “Telling More Than We Can Know: Verbal Reports on Mental Processes”, *The Psychological Review*, 84, 3, pp. 231-259.

O’Brian, L. (2003), “Moran on Agency and Self-Knowledge”, *European Journal of Philosophy*, 11 3, pp.375-390.

Peacocke, C. (1996), “Entitlement, Self-Knowledge, and Conceptual Redeployment”, *Proceedings of Aristotelian Society*, 96, pp. 177-158.

Pedrini, P. (2009), *Prima persona. Epistemologia dell’autoconoscenza*, ETS, Pisa.

Pedrini, P. “On the Failure of Authority”, manoscritto.

Perry, J. (1977), “Frege on Demonstratives”, *Philosophical Review*, 86, pp. 474-97.

Perry, J. (1979), “The Problem of Essential Indexical”, *Nous*, 13, pp. 3-21.

Searle, J. (1992), *The Rediscovery of the Mind*, MIT Press, Cambridge, Ma.

Shoemaker, S. (1963), *Self-Knowledge and Self-Identity*, Cornell University Press, Ithaca.

Shoemaker, S. (1968), “Self-Reference and Self-awareness”, *The Journal of Philosophy*, 65, 19, pp. 555-576.

Shoemaker, S. (1994a), “Self-Knowledge and ‘Inner Sense’: Lecture I: The Object-Perception Model”, *Philosophical and Phenomenological Research*, 54, 2, pp. 249-269.

Shoemaker, S. (1994b), “Self-Knowledge and ‘Inner Sense’: Lecture II: The Broad Perceptual Model”, *Philosophical and Phenomenological Research*, 54, 2, pp. 271-290.

Shoemaker, S. (1994c), “Self-Knowledge and ‘Inner Sense’: Lecture III: The Phenomenal Character of Experience”, *Philosophical and Phenomenological Research*, 54, 2, 291-314.

Stanford Encyclopaedia of Philosophy, entry “Self-Knowledge”.

Taylor, C. (1977), "Self-interpreting Animals", in Taylor (1985a)

Taylor, C. (1985b), *Philosophy and the Human Sciences: Philosophical Papers*, vol. 2, Cambridge University Press, Cambridge.

Wright, C., Smith, B. and Macdonald, C. (1998), eds., *Knowing Our Own Minds*, Oxford, Clarendon Press.

Aphex.it è un periodico elettronico, registrazione n/ ISSN 2036-9972. Il copyright degli articoli è libero. Chiunque può riprodurli. Unica condizione: mettere in evidenza che il testo riprodotto è tratto da www.aphex.it

Condizioni per riprodurre i materiali --> Tutti i materiali, i dati e le informazioni pubblicati all'interno di questo sito web sono "no copyright", nel senso che possono essere riprodotti, modificati, distribuiti, trasmessi, ripubblicati o in altro modo utilizzati, in tutto o in parte, senza il preventivo consenso di Aphex.it, a condizione che tali utilizzazioni avvengano per finalità di uso personale, studio, ricerca o comunque non commerciali e che sia citata la fonte attraverso la seguente dicitura, impressa in caratteri ben visibili: "www.aphex.it". Ove i materiali, dati o informazioni siano utilizzati in forma digitale, la citazione della fonte dovrà essere effettuata in modo da consentire un collegamento ipertestuale (link) alla home page www.aphex.it o alla pagina dalla quale i materiali, dati o informazioni sono tratti. In ogni caso, dell'avvenuta riproduzione, in forma analogica o digitale, dei materiali tratti da www.aphex.it dovrà essere data tempestiva comunicazione al seguente indirizzo (redazione@aphex.it), allegando, laddove possibile, copia elettronica dell'articolo in cui i materiali sono stati riprodotti.

In caso di citazione su materiale cartaceo è possibile citare il materiale pubblicato su Aphex.it come una rivista cartacea, indicando il numero in cui è stato pubblicato l'articolo e l'anno di pubblicazione riportato anche nell'intestazione del pdf. Esempio: Autore, *Titolo*, <<www.aphex.it>>, 1 (2010).
