

**Universidade Federal do Rio Grande do Sul
Instituto de Filosofia e Ciências Humanas
Programa de Pós-Graduação em Filosofia**

Teorias da Introspecção e psicologia moral

Eduardo Vicentini de Medeiros

Orientador: Dr. Jônadas Techio

**Dissertação apresentada como requisito parcial à obtenção
do grau de Mestre em Filosofia**

Porto Alegre, Brasil

Fevereiro de 2013

Sumário

Resumo.....	1
Abstract.....	3
Introdução.....	5
1. Definindo “Introspecção”	11
2. Modelos perceptuais da Introspecção.....	19
2.1 Olhando para dentro.....	21
2.1.1 Franz Brentano.....	21
2.1.2 William James.....	26
2.1.3 James Sully.....	43
2.1.4 David Armstrong.....	47
2.2 Olhando para fora: Gareth Evans	51
3. Modelo não-perceptual da Introspecção.....	57
3.1 Sydney Shoemaker.....	63
3.1.1 Modelo objeto-percepção e condição corpóreo- voluntária.....	66
3.1.2 Condição experiencial.....	67
3.1.3 Condição fato-objeto e condição de identificação.....	69
3.1.4 Condição não-relacional.....	74
3.1.5 Condição de atenção.....	78
3.1.6 Modelo perceptual amplo e <i>self-blindness</i>....	79

3.1.7 Constitutivismo, conjunções mooreanas e argumento moral.....	84
4. Retratos da interioridade e conceitos morais.....	93
4.1 Visão não-relacional do argumento moral.....	97
Considerações finais.....	106
Referências.....	112

Resumo

Existe alguma relação conceitual relevante entre introspecção e conceitos morais? Qual a relação entre as diferentes teorias da introspecção e temas da psicologia moral? Estas são as perguntas centrais desta dissertação. O ponto para o qual quero chamar a atenção é sobre a relação entre diferentes teorias da introspecção e teses em psicologia moral e nas teorias morais. Argumentarei a favor da seguinte afirmação: qualquer filosofia moral pressupõe determinadas teses sobre atribuição e autoatribuição de estados mentais, em especial, atribuição e autoatribuição de atitudes proposicionais. Ou seja, nenhuma teoria moral é inocente em relação a determinados pressupostos que podem ser mapeados em diferentes teorias da introspecção. Além disso, creio que o caminho inverso também mereça análise, a saber que assumir uma determinada teoria sobre o funcionamento de conceitos psicológicos pode ter implicações para o tipo de teoria moral que deva ser aceita. Nosso objetivo é demarcar dois grupos de teorias sobre o funcionamento dos processos introspectivos. De um lado as teorias que assumem uma analogia de inspiração lockiana entre sentido externo e sentido interno, quando serão analisadas teses de autores seminais da psicologia introspeccionista do final do século XIX, como Franz Brentano (*Psychologie vom empirischen Standpunkten*- 1874), William James (*The Principles of Psychology*-1890) e James Sully (*Illusions* - 1881). Examinaremos também a primeira formulação filosófica contemporânea das teorias do sentido interno, apresentada por David Armstrong (*A Materialist Theory of the Mind* - 1968).

De outro lado, discutiremos as teorias da introspecção que rejeitam o modelo perceptual. O principal responsável na filosofia contemporânea por esta crítica à analogia lockiana é Sydney Shoemaker. Faremos uma detalhada análise dos argumentos de Shoemaker apresentados nas *Royce Lectures*, de 1994. Os argumentos de Shoemaker deram origem ao que se convencionou chamar teorias racionalistas da introspecção ou do autoconhecimento, e dentre os vários

autores representativos desta tendência, escolhemos Richard Moran (*Authority and Estrangement: An Essay on Self-Knowledge* - 2001) como representativo do desenvolvimento consequente da crítica ao modelo perceptual da introspecção em direção aos temas da psicologia moral.

Argumentaremos que as teorias que assumem a analogia lockiana entre sentido externo e interno dão suporte ao que denominamos de visão não-relacional do argumento moral. Dado que os argumentos que criticam estas teorias do sentido interno são cogentes, então as características distintivas desta visão não-relacional devem ser deixadas de lado em uma análise adequada do argumento moral.

Abstract

Is there any relevant conceptual connection between introspection and moral concepts? What is the connection between different theories of introspection and issues in moral psychology? Those are the central questions of this dissertation. The point to which I wish to draw attention is the relationship between different theories of introspection and theses in moral psychology and moral theories. I shall argue in defense of the following statement: any moral philosophy presupposes certain theses on self-attribution of mental states and, in particular, self-attribution of propositional attitudes. That is, no moral theory is innocent on certain assumptions that can be mapped into different theories of introspection. Furthermore, I believe that the opposite way also deserves analysis, namely one assuming that a particular theory about the functioning of psychological concepts may have implications for the kind of moral theory that should be accepted. Our goal is to demarcate two groups of theories about the operation of introspective processes. On one side, theories that assume an analogy, inspired by Locke, between external and internal sense. In this case we will analyze theses of seminal authors associated with introspective psychology from the late nineteenth century, as Franz Brentano (*Psychologie vom empirischen Standpunkte*-1874), William James (*The Principles of Psychology*, 1890) and James Sully (*Illusions* - 1881). We will also examine the first formulation of contemporary philosophical theories of the internal sense presented by David Armstrong (*A Materialist Theory of the Mind* - 1968).

On the other side, we shall discuss the theories of introspection that reject the perceptual model. The main responsibility in contemporary philosophy for this critique of the Lockean analogy is Sydney Shoemaker. We will pursue a detailed analysis of the arguments presented in *Royce Lectures* delivered in 1994. Shoemaker's arguments gave rise to the so-called rationalist theories of introspection or self-knowledge, and among the several authors of this trend we chose Richard Moran (*Authority and Estrangement: An Essay on Self-Knowledge-*

2001) as representative of the consequent development of this critique to the perceptual model of introspection toward themes of moral psychology.

We will argue that theories that assume the Lockean analogy between internal and external senses, offer support to what we are naming non-relational view of the moral argument. Given that the arguments criticizing these theories of the internal sense are cogent, the distinctive characteristics of this non-relational view must be set aside in a proper analysis of the moral argument.

Introdução

Normalmente, as pessoas acham graça na caricatura do psicólogo behaviorista que, ao encontrar seu colega, pergunta: “Você está ótimo! E eu, como estou?”. De fato, é cômico pensar em um personagem que precisa perguntar a outra pessoa como anda seu estado de ânimo, ou mesmo olhar no espelho para saber se está feliz ou zangado. A pressuposição que justifica a estranheza e a comicidade destes casos é que, ao menos aparentemente, deveríamos ter algum tipo de conhecimento privilegiado sobre nossos estados de ânimo, o que nos leva a generalizar que deveríamos ter algum tipo de conhecimento privilegiado de outros estados mentais, como atitudes proposicionais, traços de caráter e sensações, por exemplo. Este conhecimento seria privilegiado no sentido em que a pessoa que tem ou está nestes estados teria um conhecimento direto e não-inferencial dos mesmos.

Não preciso olhar no espelho para saber que estou zangado, enquanto que, quem me observa “de fora” pode inferir a partir da expressão do meu rosto que não estou nem um pouco satisfeito com as aulas de trombone do meu vizinho. Quando observo minha esposa saindo de casa com a sombrinha a tiracolo, infiro da descrição deste comportamento que ela deveria estar pensando que irá chover hoje à tarde, que ela, possivelmente, tenha lido a mesma previsão do tempo que eu li no jornal esta manhã. Como sei, por experiências anteriores, que ela é uma pessoa bastante previdente, essa me parece uma boa explicação para seu comportamento. Agora, pode ser o caso que ela tenha levado consigo a sombrinha tão somente em razão da feliz combinação entre os tons de amarelo de sua nova calça e das listras da sombrinha. Neste caso, ela saberia disto de uma maneira completamente diferente daquela que eu poderia, mesmo que improvavelmente, saber. Ela tem um acesso direto a esta informação, enquanto que o máximo que posso fazer é supor algumas hipóteses para explicar seu comportamento. Dito de modo mais geral, a pressuposição é que deveríamos

ter um conhecimento privilegiado de tudo o que compõe nosso mobiliário mental – ou pelo menos da parte desse mobiliário que decora a sala de estar de nossa vida interior, a qual acessamos conscientemente.

Aqui, é importante demarcar claramente a área de investigação que me proponho a examinar. O que está em jogo nesta dissertação são as diferentes teorias da introspecção, tão somente na medida em que estas teorias estão interessadas em explicar o conhecimento que temos sobre nossos estados mentais conscientes. Não estarão em exame as complicadas questões epistemológicas e as ressonâncias morais da introdução do conceito de “inconsciente” tal como este conceito aparece nas teorias psicológicas de orientação psicanalítica. Bertrand Russell, quando discute em *The Analysis of Mind* a confiabilidade dos processos introspectivos, explicitamente usa as descobertas ou descrições da psicanálise sobre o funcionamento do desejo como uma objeção:

To begin with the trustworthiness of introspection. It is common among certain schools to regard the knowledge of our mental processes as incomparably more certain than our knowledge of the ‘external world’; this view is to be found in the British philosophy which descends from Hume, and is present, somewhat veiled, in Kant and his followers. There seems no reason whatever to accept this view.(...)Not only are we often unaware of entertaining a belief or desire which exist in us; we are often actually mistaken. The fallibility of introspection as regards what we desire is made evident by psycho-analysis; its fallibility as to what we know is easily demonstrated. (Russell 1921, p. 122, meu sublinhado)

Creio que aceitar este tipo de argumento é aceitar uma imagem dos processos introspectivos como sendo a única via que temos para o “conhecimento de nossos processos mentais”, onde “introspecção” seria apenas um nome para uma noção mais geral de autoexame que estaria aqui pressuposta. Espero que esta extensão equívoca do termo “introspecção” seja superada na medida em que avançarmos na direção de uma definição mais precisa do termo, o que será feito no primeiro capítulo.

Mesmo na literatura recente não estamos livres do tipo de imprecisão envolvida em achar que a introspecção deva dar conta do conhecimento que possamos ter de processos mentais não-conscientes, como lembra Schwitzgebel:

Present-day philosophers and psychologists often doubt the layperson's talent for assessing such nonconscious mental states as personality traits, motivations, skills, implicit beliefs and desires, and the bases of decisions, and they sometimes describe such doubts as doubts about 'introspection' (Schwitzgebel 2011, p.119, meu sublinhado).

Por trás da imagem vaga e um tanto imprecisa do que se convencionou chamar “autoridade de primeira pessoa”, aparecem vários apelos intuitivos sobre a relação cognitiva que temos com nossos estados mentais. Meu esforço nesta dissertação será na direção de mapear, em diferentes épocas e através de diferentes autores, algumas das teorias mais influentes sobre o funcionamento da introspecção, pois, como veremos, os temas relacionados à autoridade da primeira pessoa são intimamente ligados a determinadas imagens desse funcionamento.

Ainda mais drasticamente, criticar esta suposta autoridade de primeira pessoa pode passar pela pergunta sobre a própria possibilidade de que exista algo como a atividade de “introspectar” nossos estados mentais. Um mapeamento completo e detalhado das diferentes teorias da introspecção pressuporia percorrer um caminho bastante extenso, extenso demais para as minhas pretensões presentes. Faremos recortes nesta história mas, apesar de certa arbitrariedade na escolha pelos autores, imagino que o retrato acabará captando o essencial da discussão contemporânea sobre o tema, bem como os temas que interessam-me nesta discussão.

Este mapeamento, no entanto, é apenas metade da tarefa a que me proponho. O ponto para o qual quero chamar a atenção, em especial no capítulo “Retratos da interioridade e conceitos morais”, é a relação entre estas diferentes teorias da introspecção e teses em psicologia moral e nas teorias morais. Naquele capítulo, argumentarei a favor da seguinte afirmação: qualquer filosofia moral pressupõe determinadas teses sobre atribuição e autoatribuição

de estados mentais, em especial, atribuição e autoatribuição de atitudes proposicionais. Ou seja, nenhuma teoria moral é inocente em relação a determinados pressupostos que podem ser mapeados em diferentes teorias da introspecção. Além disso, creio que o caminho inverso também mereça análise, a saber que, assumir uma determinada teoria sobre o funcionamento de conceitos psicológicos pode ter implicações para o tipo de teoria moral que deva ser aceita.

É sintomático que alguns autores importantes na discussão de temas da filosofia da mente e da filosofia da psicologia acabem por dar relevo a este tipo de implicação. Para ficarmos apenas com dois bons exemplos, Alvin Goldman, no último capítulo de *Simulating Minds*, coloca a questão nos seguintes termos:

If mental simulation is acknowledged as a core feature of the human mind, what impact should this have, if any, on moral theory? Connections between simulation and moral theory might be charted via the closely linked notions of empathy and/or perspective taking. (Goldman 2006, p. 290)

O importante nesta citação de Goldman é o fato da sua pergunta ser formulada na forma de um condicional. Se os processos de simulação estão diretamente envolvidos numa explicação aceitável do funcionamento da mente humana, quais as implicações disto para as teorias morais? Certamente, a aceitação da simulação como uma teoria relevante no âmbito da filosofia da mente terá consequências para a disputa metaética entre sentimentalistas e racionalistas, por exemplo.

Olhando a questão por outro ângulo, na medida em que teorias éticas pressupõem um conceito de felicidade, seja esta teoria de matiz utilitarista ou outro, Goldman sugere que a simulação possa servir como uma ferramenta epistemológica para atribuição de estados de felicidade a outros ou mesmo na produção de experiências de felicidade através dos mecanismos associados de contágio emocional. Estas seriam três aplicações possíveis de uma teoria da simulação no domínio da discussão moral.

No último capítulo de seu mais recente livro, *The Opacity of Mind*, Peter Carruthers defende uma teoria integrativa do autoconhecimento que denomina ISA (*Interpretative Sensory-Access*), e também coloca a questão na forma de um condicional:

If there are no conscious decisions then it seems to follow that there is no such thing as conscious agency. And if there are no conscious agency, then there are no conscious agents. (At least, this is so in one good sense of the phrase ‘conscious agent.’) For one might think that a conscious agent would be an agent who arrives at *conscious* decisions, and who are capable of exercising conscious control over choice and action. And surely conscious *agency* requires such control. But if ISA theory is true,(...) then it seems to follow that no such thing exists. For it turns out that there are no conscious decisions. (Carruthers 2011, p.381)

Neste caso, parece óbvio que se aceite uma teoria no âmbito da filosofia da mente que me leva a negar a existência de decisões conscientes, então sou obrigado a repensar de modo radical o conceito de responsabilidade tal como apresenta-se nas teorias morais. E é esta implicação que Carruthers corretamente indica como uma consequência de sua teoria.

Esta introdução não é o lugar para discutir as teorias de Goldman ou Carruthers. Aliás, uma discussão como esta pressuporia o tratamento prévio de um sem número de temas que não estão em nosso foco. Meu ponto é apenas mostrar que a relação entre estes domínios de investigação – filosofia da mente e filosofia da psicologia por um lado, e psicologia moral e teoria moral por outro – não é acidental.

Em um artigo que marcou época, Elisabeth Anscombe traçou um desafio para a filosofia moral, enunciada na primeira tese defendida no artigo, que afirma o seguinte:

... it is not profitable for us at present to do moral philosophy; that should be laid aside at any rate until we have an adequate philosophy of psychology, in which we are conspicuously lacking (Anscombe 1958, em “Modern Moral Philosophy”, publicado originalmente na revista *Philosophy* 33/ 124)

Parte do trabalho de construir uma “adequada filosofia da psicologia” foi desenvolvido pela própria Anscombe em *Intention*, de 1957, mas não creio que possamos com tranquilidade dizer que temos algum tipo de consenso entre os filósofos, de 1957 até nossos dias, sobre qual a interpretação adequada para os verbos psicológicos que especialmente interessam para o domínio da moral, verbos como “intencionar”, “desejar”, “querer” e outros assemelhados, ou mesmo algum consenso sobre o modo como conhecemos nossos estados mentais ou os estados mentais de outras pessoas. Por mais que não tenhamos dado todos os passos na direção correta, e estejamos marcando passo em um terreno não muito distante daquele descrito por Anscombe, creio que é completamente defensável afirmar que já temos bons argumentos para indicar onde não devemos ir. Em particular, creio que temos bons argumentos para deixar de lado as teorias perceptuais da introspecção. E o principal responsável por este resultado é Sydney Shoemaker. No capítulo 3.1 daremos especial atenção ao exame das teses de Shoemaker tal como foram defendidas nas *Royce Lectures*, de 1994, e quais seriam os desdobramentos destes resultados para a psicologia moral.

Capítulo 1 - Definindo “Introspecção”

Não há nada de especial ou extraordinário na experiência de “introspectar” nossos estados mentais. Num certo sentido, podemos afirmar que “pensar sobre algo” e “introspectar” são a mesma atividade. Vejamos até onde esta analogia nos leva. Você estaria disposto a considerar um relato introspectivo algo do tipo: “Estou pensando na árvore da esquina”? A primeira reação é pensar que os objetos da introspecção são estados internos como desejos, intenções ou estados de ânimo como raiva ou esperança. Certamente, a árvore da esquina não é um estado interno, mas o pensamento sobre a árvore da esquina é um estado interno. É neste sentido também que falamos não apenas do pensamento sobre a árvore da esquina, mas também podemos falar de “introspectar” nossa experiência sensorial da árvore da esquina, seja esta experiência visual, auditiva, olfativa, tátil ou gustativa.

Um ponto que está além de qualquer controvérsia é que não “introspectamos” estados internos de outras pessoas. “Introspectar” é uma atividade, por definição, conduzida do ponto de vista da primeira pessoa. Se fôssemos equipados com a extraordinária capacidade observar diretamente os estados mentais de outras pessoas, seria impossível, ou pelo menos altamente improvável para a filosofia, desenvolver ficções de acesso privilegiado ou dramatizar o ceticismo sobre outras mentes. Mas, infelizmente (ou felizmente como me parece mais sensato pensar, pois um mundo de pessoas transparentes neste sentido seria um pesadelo moral), não temos esta capacidade.

Alguns filósofos, como Gilbert Ryle em *The Concept of Mind* (1949), chegaram a pensar que o procedimento pelo qual atribuímos estados mentais a nós mesmos é exatamente o mesmo que usamos para atribuir estados mentais a outros, ou seja, que não há tal coisa como a introspecção, considerada como

uma atividade peculiar ao ponto de vista da primeira pessoa e que tenha uma epistemologia também peculiar.

Outro ponto que parece incontroverso sobre as características essenciais dos processos ditos introspectivos é que eles nos dão informação sobre estados ou processos mentais, tão somente. O conhecimento que tenho sobre a posição atual de minhas pernas, por exemplo, mesmo que possamos caracterizar como um conhecimento não-observacional, não é um relato introspectivo. Neste caso, parece que estamos perto de um limite um tanto difuso, mas certamente o conhecimento que tenho sobre meu peso ou minha altura, ou até mesmo sobre a roupa que estou usando, não são exemplos de conhecimentos que tenho por introspecção.

Entre as características definitórias dos processos introspectivos podemos ainda incluir mais um traço: os processos introspectivos nos dão conhecimento de nossos estados mentais ocorrentes, atuais ou imediatamente passados. É no mínimo duvidoso que possamos descrever como introspectivo um relato atual que descreva as razões que me fizeram escolher estudar filosofia quando me matriculei no vestibular de 1987. Igualmente duvidoso dizer que é por introspecção que posso afirmar que ficaria muito feliz se ganhasse um DVD do Black Label Society no próximo Natal. Sejam quais forem os processos envolvidos nestes casos, a introspecção não parecer estar entre eles.

Estas três características correspondem ao que Eric Schwitzgebel nomeia, respectivamente, condição de primeira pessoa, condição de mentalidade (*mentality condition*) e condição de proximidade temporal no verbete *Introspection* para a versão online da Stanford Encyclopedia of Philosophy (<http://plato.stanford.edu/entries/introspection/>)

Para um exame mais detalhado, nosso ponto de partida não poderia ser outro do que nos entendermos sobre o que exatamente estamos querendo dizer quando usamos o termo “introspecção”. Para isso, vejamos algumas definições.

Dicionário conciso de Oxford (OED):

Introspect v.i. examine one's own thoughts and feelings; hence ~ION n., examination or observation of one's own mental processes

James Sully, *Illusions* – 1887:

This inquiry naturally sets out with the question: What is meant by introspection? This cannot be better defined, perhaps, than by saying that it is the mind's immediate reflective cognition of its own states as such. (p. 64)

William James, *The Principles* – 1890:

The word introspection need hardly be defined – it means, of course, the looking into our own minds and reporting what we there discover. *Everyone agrees that we there discover states of consciousness* (p. 185)

Bertrand Russell, *The Analysis of Mind* – 1921:

When you see a friend coming along the street, you acquire knowledge of an external, physical fact; when you realize that you are glad to meet him, you acquire knowledge of a mental fact. Your dreams and memories and thoughts, of which you are often conscious, are mental facts, and the process by which you become aware of them *seems* to be different from sensations. Kant calls it the 'inner sense'; sometimes it is spoken as 'consciousness of self'; but its commonest name in English psychology is 'introspection'. (p. 109)

David Armstrong, *A Materialist Theory of the Mind* – 1968:

In sense-perception we become aware of current happenings in the physical world. A perception is therefore a mental event having as its (intentional) object situations in the physical world. In introspection, on the contrary, we become aware of current happenings in our own mind. Introspection is therefore a mental event having as its (intentional) object other mental happenings that form part of the same mind. (p. 323)

Sydney Shoemaker, *Self-knowledge and Inner Sense* – 1996:

Our minds are so constituted, or our brains are so wired, that, for a wide range of mental states; one's being in a certain mental state produces in one, under certain

conditions, the belief that one is in that mental state. This is what our introspective access to our mental states consists in. (p. 222)

Eric Schwitzgebel, *Introspection, What?*- 2012:

I am inclined to recommend the following view: introspection is the dedication of central cognitive resources, or attention, to the task of arriving at a judgment about one's current, or very recently past, conscious experience, using or attempting to use some capacities that are unique to first-person case (...), with the aim or intention that one's judgment reflect some relatively direct sensitivity to the target state" (p.42-3)

Em todas estas definições temos a satisfação da condição de mentalidade, seja falando do exame de nossos pensamentos e sentimentos, seja na observação de nossos próprios processos mentais, cognição reflexiva imediata dos estados da mente, estados da consciência ou eventos mentais ocorrentes.

Apenas nas definições de Schwitzgebel e James fica claro que a introspecção trata de eventos ou processos mentais conscientes. No entanto, falar do acesso aos conteúdos mentais conscientes ainda é muito vago. Cabe muita coisa dentro deste rótulo. Os candidatos mais naturais para mobiliar a sala de estar de nossa mente são, de acordo com a classificação proposta por André Gallois em *The World Without, the Mind Within* (1996), os seguintes:

- (a) estados de ânimo, tais como depressão, euforia, ansiedade e desespero;
- (b) traços de caráter tais como ser preguiçoso, generoso, alegre, corrupto, cruel, insensível, gentil, tolerante ou perspicaz;
- (c) emoções como amor, aversão, raiva e medo;
- (d) atitudes proposicionais como acreditar, temer, intencionar, desejar, esperar, saber, arrepende-se, perceber, compreender;
- (e) sensações e experiências (visual, auditiva, tátil, etc)

(Gallois 1996, p. 17)

Não parece sensato pensar que, para um conjunto tão díspar, seja possível um tratamento uniforme. Esta suspeita levará água para o moinho de teorias da introspecção que a definam como uma pluralidade de processos, como está previsto na definição de Schwitzgebel acima. Quando ele fala, no plural, de “recursos cognitivos centrais” ou “capacidades que são próprias da posição da primeira pessoa”, já podemos antever que, quando falamos de introspecção, podemos estar falando da confluência simultânea de uma variedade de processos cognitivos que se combinam de maneiras diferentes, dependendo do tipo de processo ou estado mental a ser introspectado.

Se eu fizer uma lista dos meus estados mentais ocorrentes neste momento, ficará claro, ao menos para mim, que a introspecção não é um processo que possa ter seu funcionamento explicado de maneira uniforme para todos os casos. A multiplicidade de tipos de estados mentais sugere que a introspecção ocorre de diferentes maneiras. Se você tiver dificuldade em aceitar o caráter múltiplo dos aspectos introspectivos, procure descrever seus conteúdos mentais ocorrentes tal como a lista que apresento a seguir.

Minha predisposição para jogar xadrez, a estafa ao final de um dia de estudos, o desejo por uma banana split de creme, morango e abacate, a lembrança do aroma do cappuccino que experimentei na semana passada, a saudade de todos os cachorros que já tive, em especial quando acompanhada da imagem detalhada de cada um deles, a grata satisfação de ter preparado à perfeição o nhoque do último almoço, o remorso por ter exagerado neste mesmo almoço, a sensação de alguns quilos a mais, trechos do Ulisses, imagens de algumas telas de van Gogh, melodias gregorianas, a preguiça de escrever a esta hora da noite, a inquietação para terminar esta lista, o frio nas mãos, a recordação de que me recomendaram usar luvas de lã, a consciência de que eu já deveria estar dormindo pois amanhã cedo tenho de tirar o lixo da área de serviço, lavar a louça, molhar as plantas, ir ao supermercado comprar mais vinho e algumas frutas; o descontentamento com o desempenho de meu time, o medo derivado de um princípio de azia e de uma leve pressão no olho

esquerdo etc, etc. Todas essas coisas eu sei por introspecção. Estão presentes à minha consciência no momento em que escrevo estas linhas. Sua diversidade é tão manifesta que me faz duvidar que o processo pelo qual sou delas consciente seja um e somente um tipo de processo.

De um modo bastante resumido, podemos separar as diferentes teorias da introspecção em dois grandes grupos. De um lado, as teorias perceptuais que desenvolvem de diferentes maneiras a analogia básica entre sentido externo e sentido interno. Esta analogia tem sua expressão canônica em *An Essay Concerning Human Understanding* de John Locke:

“...the other fountain from which experience furnisheth the understanding with ideas is, – the perception of the operations of our own mind within us, as it is employed about the ideas it has got; – which operations, when the soul comes to reflect on and consider, do furnish the understanding with another set of ideas, which could not be had from things without. And such are perception, thinking, doubting, believing, reasoning, knowing, willing, and all the different actings of our own minds; – which we being conscious of, and observing in ourselves, do from these receive into our understandings as distinct ideas as we do from bodies affecting our senses. This source of ideas every man has wholly in himself; and though it be not sense, as having nothing to do with external objects, yet it is very like it, and might properly enough be called internal sense.” (Locke 1689/1975, II, 1.iv)

Na epistemologia lockiana esta analogia entre sentido externo e interno cumpre a função de explicar a origem das ideias ou a partir da sensação ou da reflexão. As expressões que Locke utiliza para caracterizar o funcionamento do sentido interno moldaram a discussão sobre o tema não apenas no cenário da filosofia empirista. Os reflexos desta análise podem ser vistos até mesmo no desenvolvimento da psicologia experimental do final do século XIX. As citações de James e Sully retratam o ponto. James fala de “olhar para dentro de nossa mente” para “descobrir” estados de consciência. Sully usa a expressão mais resguardada, mas de forte matiz lockiano, de uma “cognição reflexiva imediata”. Creio, inclusive, que a analogia lockiana moldou as opiniões de senso comum sobre o tema, dado que a definição do OED citada acima, usa

exatamente os mesmos termos (“examinar”, “observar”). Quando Locke fala de “percepções das operações de nossa própria mente”, ou que podemos estar conscientes das ações de nossa mente “observando em nós mesmos”, ele está formatando a ortodoxia do tratamento da introspecção como algo passivo, algo que ocorre ao sujeito. Existe um estado mental prévio de crença, desejo ou dúvida e o sujeito percebe ou observa a ocorrência deste estado, ou nos termos de James, “descobrimos” a ocorrência deste estado. Ou, nos termos de Armstrong, nos tornamos conscientes de que um estado de crença, desejo ou dúvida acontece em nossa mente.

Seguindo com nossa classificação em dois grandes grupos, do outro lado temos as teorias não-perceptuais da introspecção, cuja principal característica é a recusa da analogia entre sentido externo e interno como relevante para a explicação do funcionamento dos processos pelos quais conhecemos nossos estados mentais. Na definição de Shoemaker citada acima, a relação entre o ato de introspectar e o estado mental introspectado não é caracterizada como algo que eu observo/percebo *versus* algo que é observado/percebido: “one’s being in a certain mental state produces in one, under certain conditions, the belief that one is in that mental state”. Que os estados mentais de um sujeito sejam acessíveis introspectivamente não é um traço contingente, causalmente relacionado. Ter um desejo e crer que se tem um desejo é uma relação constitutiva, não-contingente. E o mesmo vale para qualquer estado mental consciente: tenho um estado mental consciente x se, e somente se, creio que tenho um estado mental consciente x. No subcapítulo 3.1.7 que examina as teses de Shoemaker, exploraremos algumas consequências para a psicologia moral desta caracterização.

Nosso interesse nesta dissertação está focado na relação entre as diferentes teorias da introspecção e a relação destas teorias com os temas da psicologia moral. Sendo assim, o conjunto de estados mentais formado pelos traços de caráter, como a preguiça, intolerância ou insensibilidade, deveria nos interessar diretamente. Mas aqui aparece um primeiro problema. Se definirmos a introspecção de acordo com a condição da proximidade temporal, e se

definirmos traços de caráter como disposições, parece que temos um obstáculo ao conhecimento introspectivo dos traços de caráter. A disposicionalidade da minha preguiça, por exemplo, é expressa condicionalmente: se tais e tais condições forem o caso (procrastinar continuamente decisões, perder prazos e acumular tarefas não resolvidas podem ser algumas destas condições), então meu estado pode ser descrito como preguiça. A satisfação das condições elencadas demanda desenvolvimento continuado no tempo, o que nos leva a concluir que traços de caráter são não-episódicos. Sendo assim, a introspecção sozinha, dado que seus resultados são sempre relatos de estados mentais ocorrentes ou muito próximos no passado, não é condição suficiente (sem ajuda da memória e imaginação) para autoatribuição de traços de caráter, definidos disposicionalmente. Se este argumento tem alguma força, devemos abrir mão ou da condição de proximidade temporal, ou da caracterização disposicional dos traços de caráter.

Capítulo 2 – Modelos Perceptuais da Introspecção

O que propomos neste capítulo é um recorte na história das teorias da introspecção, que começa com o surgimento da psicologia introspeccionista do final do século XIX, em especial com a publicação de *Psychologie vom empirischen Standpunkte*, de Franz Brentano, em 1874. Brentano (1838 - 1917), William James (1842 - 1910) e James Sully (1842 - 1923) são as figuras principais que compõem nosso retrato dos primeiros impulsos da psicologia na direção de firmar-se como uma ciência autônoma, com métodos e objetos de estudo definidos e que tem nos processos introspectivos seu principal cavalo de batalha. Todos estes autores estão comprometidos, de uma maneira ou outra, com uma concepção da introspecção que tem suas origens na definição lockiana de sentido interno. Sendo assim, todas as teorias do sentido interno sobre a introspecção assumem que a introspecção guarda uma similaridade com a percepção externa. Vários autores apresentaram roupagens contemporâneas para esta intuição lockiana fundamental. William Lycan, em *Consciousness and Experience* (1996), Shaun Nichols e Stephen Stich em *Mindreading* (2003) e Alvin Goldman em *Simulating Minds* (2006) são alguns exemplos. No entanto, examinaremos apenas a primeira formulação contemporânea de uma teoria do sentido interno da introspecção oferecida por David Armstrong em *A Materialist Theory of Mind*, publicado em 1968 e revisado em 1993. O exame destes autores - Franz Brentano, William James, James Sully e David Armstrong - será suficiente para termos uma ideia do funcionamento das teorias do sentido interno da introspecção, que serão detalhadas neste capítulo.

Ainda dentro do marco das teorias do sentido interno, mas apontando em outra direção, analisaremos a noção de “transparência” desenvolvida por Gareth Evans. Quando queremos explicar a autoatribuição de estados mentais tal como expresso em sentenças como “Creio que vai chover amanhã”, “Penso que hoje é quarta-feira” (autoatribuição de atitudes proposicionais), “Deste ângulo A parece maior que B” (autoatribuição de uma experiência visual), uma

outra analogia entre introspecção e percepção externa vem à tona: talvez possamos explicar algumas de nossas alegações de autoconhecimento através do processo de olhar para fora, para o mundo, para o modo como as coisas exteriores a nós (para usar uma expressão kantiana) nos aparecem. Explorar esta analogia foi a realização de Gareth Evans em seu capítulo sobre autoidentificação em *The Varieties of Reference* (1982)

2.1 Olhando para dentro: Franz Brentano, William James, James Sully e David Armstrong

2.1.1 Franz Brentano

Franz Brentano me interessa pela distinção que ele introduziu em sua obra *Psychologie vom empirischen Standpunkte*, de 1874, entre “percepção interna” (Inner Perception/*Wahrnehmung*) e “observação interna” (Inner Observation/*Beobachtung*). Esta distinção aparece no começo do segundo capítulo intitulado “Psychological Method with Special Reference to its Experiential Basis”, que passo a citar na tradução inglesa:

Psychology, like the natural sciences, has its basis in perception and experience. Above all, however, its source is to be found in the *inner perception* of our own mental phenomena. We would never know what a thought is, or a judgment, pleasure or pain, desires or aversions, hopes or fears, courage or despair, decisions and voluntary intentions if we did not learn what they are through inner perception of our own phenomena. Note, however, that we said that inner *perception* [*Wahrnehmung*] and not introspection, i.e. inner *observation* [*Beobachtung*], constitutes this primary and essential source of psychology. These two concepts must be distinguished from one another. (Brentano 1874, p. 22)

Esta distinção é de difícil compreensão e parece artificial. Ecoa um dos principais argumentos de Comte contra a introspecção presente em *Cours de Philosophie Positive* de 1830 – que Brentano cita um pouco mais adiante no seu texto tomado da reedição de 1864, onde compra na íntegra o argumento – sobre a impossibilidade que a consciência se divida entre (A1) o ato de observar internamente um determinado fenômeno mental e (A2) o ato de estar neste estado mental a ser observado. O primeiro exemplo que Brentano usa é a raiva:

In observation, we direct our full attention to a phenomenon in order to apprehend it accurately. But with objects of inner perception this is absolutely impossible. This is especially clear with regard to certain mental phenomena such as anger. If someone is in a state in which he wants to observe his own anger raging within him, the anger must already be somewhat diminished, and so his original object of observation would have disappeared. The same impossibility is also present in all other cases. (Brentano 1874, p. 22)

Os estados de “querer observar a raiva” (A1, t1) e “estar com raiva” (A2, t1) não podem coexistir na consciência ao mesmo tempo. O argumento exposto com este exemplo supõe que, para que o sujeito queira observar a raiva, para que A1 seja o caso, A2 “must be already somewhat diminished.” Querer observar a raiva implica direcionar a **atenção** (e este é conceito fundamental na explicação de Brentano) para o “estado-de-estar-com-raiva”. E, ao fazê-lo, algo da raiva (A2 em t0) se perde ou deixa ser o que era em t0.

Brentano expressa esta impossibilidade na forma de uma lei:

It is a universally valid psychological law that we can never focus our *attention* upon the object of inner perception. (Brentano 1874, p. 22)

É revelador que o fenômeno da atenção seja o ponto relevante na formulação do argumento de Brentano. Como veremos mais adiante, ter uma compreensão adequada do que se passa quando atentamos para um estado ou fenômeno mental é um dos passos fundamentais numa teoria satisfatória da introspecção. Num primeiro momento, o que parece estar por trás desta lei que Brentano formula é o que vou chamar de princípio da não-composicionalidade ou sincronicidade dos estados mentais, ou seja, diferentes estados mentais não podem existir simultaneamente. Brentano tenta responder a esta questão quando trata da distinção entre fenômenos físicos e mentais, colocando-se a seguinte questão:

Whether and in what sense it is correct to say that mental phenomena only exist one after the other, while numerous physical phenomena can exist at the same time.

(Brentano 1874, p. 72)

William Lyons, ao tentar explicar esta distinção, coloca as coisas nos seguintes termos:

We could not, he said, directly focus on or observe our inner mental life because, in doing so, we would thereby draw away the attention necessary for the existence of the first-order mental life of thoughts, feelings and volitions. To attempt direct inner observation therefore was ipso facto to diminish or destroy what one was attempting to observe. The best one could hope to do was to notice or perceive indirectly, out of the corner of one's "mental eye", mental phenomena as they went about their business. (...) One did not actively focus on mental events – mental events by their very nature forced themselves into our notice – and this was the only way of ever finding out about them.

(Lyons 1986, p. 4)

Dizer que o melhor que podemos fazer é “perceber indiretamente” os fenômenos mentais não ajuda muito a tornar a distinção de Brentano palatável. Lyons está apenas parafraseando o que o próprio Brentano escreveu:

It is only while our attention is turned toward a different object that we are able to perceive, incidentally, the mental processes which are directed toward that object.

(Brentano 1874, p. 22).

Perceber fenômenos mentais indireta ou incidentalmente é a melhor descrição que temos até agora do que seria “percepção interna.” De fato, não é nada animador. Fico com a suspeita de que a distinção entre “percepção interna” e introspecção proposta por Brentano é, na verdade, uma solução *ad hoc* para escapar a uma objeção formulada por Comte. Um passo adiante para uma melhor compreensão talvez fosse admitir, como faz Peter Simons, que “percepção interna” é observação pela memória de um determinado fenômeno mental:

Brentano counters by distinguishing inner observation, which is impossible, from inner perception, which is the incidental awareness of all my mental phenomena concurrently with being aware of their objects. This inner perception, which Brentano, following Descartes, regards as infallible, is, assisted by memory, the experiential basis of all psychology. (Brentano 1874, introdução da segunda edição, p. XV)

Mas se interpretarmos “percepção interna” desta maneira, ficamos com o problema de compatibilizar a pretensa infalibilidade da “percepção interna” com os duvidosos mecanismos da memória, o que é identificado pelo próprio Brentano:

To be sure, this procedure, which we could call observation in memory, is obviously not fully equivalent to genuine observation of present events. As everyone knows, memory is, to a great extent, subject to illusion, while inner perception is infallible and does not admit of doubt. When the phenomena which are retained by the memory are substituted for those of inner perception, they introduce uncertainty and the possibility of many sorts of self-deception into this area at the same time. (Brentano 1874, p. 26)

A “percepção interna” não é o mesmo ato do que a observação pela memória, são modos distintos de obter dados sobre os fenômenos mentais: seja pela “percepção interna”, pela observação através da memória ou pela simples observação externa de expressões da vida mental de outras pessoas, quando, por exemplo, alguém ruboriza quando sente-se envergonhado. Isso fica claro na seguinte passagem, que finaliza o segundo capítulo “Psychological Method with Special Reference to its Experiential Basis”:

What we have said is sufficient to show from which areas the psychologist gains the experiences upon which he bases his investigation of mental laws. We found *inner perception* to be his primary source, but it has the disadvantage that it can never become observation. To inner perception we added the contemplation of our previous mental experiences in *memory*, and in this case it is possible to focus attention on them and, so to speak, observe them. The field of experience which up to this point is limited to our own mental phenomena was then extended, in that *expressions* of the mental life of other persons allow us to gain some knowledge of mental phenomena which we do not

experience directly. Certainly the facts which are important for psychology are thus increased a thousandfold. This last type of experience, however, presupposes observation through memory, just as the latter presupposes the inner perception of present mental phenomena. Inner perception, therefore, constitutes the ultimate and indispensable precondition of the other two sources of knowledge. Consequently, and on this point traditional psychology is correct as against Comte, inner perception constitutes the very foundation upon which the science of psychology is erected.

(Brentano 1874, p. 32)

Mesmo que Brentano não deixe claro que tipo de estado mental é a “percepção interna”, sabemos que a observação dos fenômenos mentais através da memória é a forma de tratar experimentalmente estas “percepções internas”. Se a psicologia ficasse restrita apenas às “percepções internas”, o sarcasmo de James atingiria por inteiro a metodologia de Brentano, como sugere Boring:

Brentano wrote in 1874: "The phenomena inwardly apprehended are true in themselves. As they appear . . . so they are in reality. Who then can deny that in this a great superiority of psychology over the physical sciences comes to light?" Against this view, James remarked: "If to *have* feelings or thoughts in their immediacy were enough, babies in the cradle would be psychologists, and infallible ones" (Boring 1953, p. 171).

2.1.2 William James

Aquelas pessoas que, por alguma razão, se interessam pela biografia de Wittgenstein, sabem do apreço que ele sempre dedicou a *The Varieties of Religious Experience*, que William James publicou em 1902, a ponto de recomendar fortemente a leitura a seus amigos, bem como dedicar-se à releitura da obra como afirma em uma carta a Russell: “Whenever I have time now I read James’s *Varieties of Religious Experience*. This book does me a lot of good”.¹

Por outro lado, ficamos com a imagem de que Wittgenstein considerava *The Principles of Psychology* uma espécie de rico repositório de incongruências e vertente da qual jorravam problemas filosóficos aos borbotões. Por trás e ao lado das declarações do credo de uma epistemologia positivista, *The Principles* é rico em metafísica. Sabemos, através da correspondência com Rush Rhess, que Wittgenstein chegou a pensar em dar um curso em Cambridge, quando do final da II Guerra, usando *The Principles* como referência. E, certamente, este curso não seria um elogio a James.

Uma forma clara e direta de mostrar o interesse contínuo que Wittgenstein demonstrou em relação aos *The Principles* é notar o paralelismo entre as passagens dos dois autores. No segundo apêndice de *A Wittgenstein Workbook*, os autores relacionam 21 temas diferentes, nos quais os textos de Zettel, *Philosophical Investigations* e *The Principles of Psychology*, se cruzam. Neste cruzamento encontramos temas decisivos para o desenvolvimento da filosofia madura de Wittgenstein como “ação intencional”, “Self” e “relação entre pensamento e linguagem” e não seria nada difícil traçar paralelos com *The Brown Book* ou *The Big Typescript*. O trabalho de mostrar toda rica tecitura de relações entre James e Wittgenstein não é nossa preocupação aqui. Este trabalho já está delineado, por exemplo, de maneira sistemática em *Wittgenstein e William James* (2004), de Russell B. Goodman.

¹ *Letters to Russell, Keynes, and Moore*, ed. G. H. von Wright (Oxford: Basil Blackwell, 1974)

De todos os nomes da psicologia introspeccionista do começo do século XIX, James certamente não foi o mais profícuo no uso sistemático da observação introspectiva sob condições, digamos, ‘laboratoriais’. Certamente, quem mais se empenhou nesta direção foi Edward Titchener, como pode ser visto no artigo “*Titchener’s Introspective Training Manual*”, em *Perplexities of Consciousness* (2011), de Eric Schwitzgebel. No entanto, considerando em retrospectiva, Franz Brentano, Wilhelm Wundt, William James e Edward Titchener, não seria nenhuma injustiça afirmar que James foi o mais denso filosoficamente entre estes autores. A experiência de ler *The Principles of Psychology* é cheia de descobertas impactantes. A riqueza e originalidade das descrições introspectivas de James são de tirar o fôlego e possuem a propriedade incomum de despertar não apenas curiosidade mas, especialmente, perplexidade. Algo do tipo: “Mas como nunca me dei conta disso!”. Várias das descrições introspectivas de James têm a feliz propriedade de trazer à mente cenas que sempre estiveram frente aos olhos, mas não nos dávamos ao trabalho de perceber com a intensidade devida.

Sem pretender generalizações, esta parece ser uma característica do espanto que a filosofia provoca quando nos dá uma perspectiva completamente diferenciada sobre assuntos que rondam nossa intimidade, nosso cotidiano.

As palavras “introspecção”, “introspectivo” e “introspectivamente” têm 76 ocorrências em *The Principles*, aparecendo praticamente durante todo percurso do livro. A maioria destas ocorrências – 19 –, como seria de se esperar, está no capítulo metodológico “The Methods and Snares of Psychology”. As outras passagens estão distribuídas pelos capítulos do livro da seguinte maneira:

The Functions of the Brain: 1

The Mind-Stuff Theory: 2

The Stream of Thought: 11

The Consciousness of Self: 7

Attention: 1

Conception: 3
Discrimination and Comparison: 5
Association: 1
The Perception of Time: 1
Memory: 2
Imagination: 3
The Perception of “Things”: 1
The Perception of Space: 2
Instinct: 1
The Emotions: 5
Will: 11

Não é necessário fazer uma análise, mesmo que não exaustiva, destas diferentes ocorrências para mapearmos que sentido (ou sentidos) James atribui ao termo “introspecção” e seus correlatos. Seu uso é bastante uniforme. Seria um exagero, no entanto, dizer que James tenha apresentado uma teoria da introspecção, mas para realçar alguns pontos de seu enfoque começarei por aquela que é, provavelmente, a mais citada dentre as frase de James quando se trata de seus pressupostos metodológicos ou de sua profissão de fé em defesa da psicologia introspectiva:

*Introspective Observation is what we have to rely on first and foremost and always. The word introspection need hardly be defined –it means, of course, the looking into our own minds and reporting what we there discover. Everyone agrees that we there discover states of consciousness. So far as I know, the existence of such states has never been doubted by any critic, however sceptical in other respects he may have been. That we have cogitations of some sort is the *inconcussum* in a world most of whose other facts have at some time tottered in the breath of philosophic doubt. All people unhesitatingly believe that they feel themselves thinking, and that they distinguish the mental state as an inward activity or passion, from all the objects with which it may cognitively deal. I regard this belief as the most fundamental of all postulates of Psychology, and shall discard all curious inquiries about its certainty as too metaphysical for the scope of this book. (James 1890, p. 185)*

Este trecho é citado por vários dos comentadores de James. Aparece em William Lyons (1986, p. 7) como evidência da filiação de James ao empirismo, tanto na versão filosófica de origem britânica de John Stuart Mill como na releitura dos psicólogos empiristas escoceses como James Mill e Alexander Bain. Lyons também considera a defesa de James dos métodos introspectivos na psicologia um resultado esperado, dado seu compromisso com uma posição dualista sobre o problema da relação mente-corpo, distinto do monismo neutro que ele desenvolveu em escritos posteriores ao *The Principles* e endossado por Russell em *The Analysis of Mind*.

Edwin G. Boring, em seu clássico artigo de 1953, "*A History of Introspection*", cita a mesma passagem dando relevo à defesa cartesiana que James faz da existência indubitável estados de consciência – *cogitationes* no jargão cartesiano – como o objeto de estudo da psicologia. Esta defesa de James fazia eco da posição comum à época. De acordo com Boring:

In general the philosophers, physiologists, and physicists who founded the new experimental psychology in 1850-1870—Fechner, Lotze, Helmholtz, Wundt, Hering, Mach, and their associates—were psychophysical parallelists who would have subscribed to this view of James. Psychology—even the new "physiological psychology"—was essentially the study of consciousness, and its chief method was introspection. (Boring 1953,p. 170)

Titchener também cita esta passagem de James em seu artigo "*Prolegomena to a Study of Introspection*" publicado em 1912, onde passa em revista "opinions of the leading psychologists of our time" (Titchener 1912, p. 428) em defesa do método introspectivo.

Eric Schwitzgebel recorre a esta passagem-chave para situar o lugar de James no começo da psicologia experimental da segunda metade do século XIX (Schwitzgebel 2011, p. 72). Se justapormos a esta citação recorrente de James partes do que vem logo depois no texto do *The Principles*, em especial esta:

But, since the rest of this volume will be little more than a collection of illustrations of the difficulty of discovering by direct introspection exactly what our feelings and their relations are, we need not to anticipate our own future details, but just state our general conclusion that *introspection is difficult and fallible; and that difficulty is simply that of all observation of whatever kind.*" (James 1890, p. 191)

então teremos encontrado um caso de um grande pensador que

...unambiguously hold, as I do, that introspection of current conscious experience is both (i) possible, important, and central to the development of a full scientific understanding of the mind and (ii) highly untrustworthy, at least as commonly practiced. (Schwitzgebel 2011, p. 118)

Gerald E. Myers, em seu artigo "Pragmatism and Introspective Psychology" para *The Cambridge Companion to William James*, usa esta passagem para introduzir a discussão sobre uma "curiosa ambiguidade" (p. 12) no uso que James faz do termo "introspecção": de um lado como retrospectção, de outro como um tipo de observação interna. Dificuldades teóricas no uso deste conceito, por parte de James, foram apontadas por William Lyons em "*The Disappearance of Introspection*", mas para explicar a relação importante entre introspecção e retrospectção, gostaria de comentar a citação destacada acima em seu contexto, salientando sua função no capítulo em que está inserida: "The Methods and Snares of Psychology", onde aparece sob o seguinte subtítulo - "The Methods of Investigation" -, o qual, por sua vez, tem as seguintes subdivisões, cujo conteúdo passo a reconstruir, comentando:

- *Introspective Observation* - todo o texto desta subdivisão é a citação que apresentamos acima;

- *A Question of Nomenclature* - neste trecho são discutidas as opções terminológicas, disponíveis até a época de James, para a expressão "states of consciousness". Segue-se do texto de James que podemos definir introspecção como o processo de olhar dentro de nossas mentes e relatar estados de

consciência. Ao explorar as opções para descrever o que exatamente é relatado ou descoberto pelo processo introspectivo, James acaba por recomendar um uso disciplinado de “pensamento/thought” e “sentimento/feeling”, descartando como menos interessantes, por variadas razões, expressões como “afecções da alma”, “ideias”, “estado mental”, “modificações conscientes”, “condição subjetiva”, e “modificações do ego”.

- *The inaccuracy of introspective observation* - Nesta passagem, James discute as objeções feitas ao método introspectivo em psicologia, já clássicas à época, que Comte apresentou em *Cours de Philosophie Positive*. James contrasta a posição comteana com o otimismo de Franz Brentano, Friedrich Ueberweg e J. Mohr, comprando o argumento de John Stuart Mill contra Comte. A objeção de Comte pode ser apresentada, grosseiramente, da seguinte maneira: a observação imediata de estados internos, como pressuposta pela psicologia introspeccionista, não é possível pois pressupõe uma divisão do sujeito entre o ato de perceber e o estado percebido que é a priori inconcebível:

But, as for observing in the same way *intellectual* phenomena at the time of their actual presence, that is a manifest impossibility. The thinker cannot divide himself into two, of whom one reasons whilst the other observes him reason. The organ observed and the organ observing being, in this case, identical, how could observation take place? This pretended psychological method is then radically null and void.” (apud James 1890, p. 188).

James responde a esta objeção citando a obra que John Stuart Mill publicou em 1865:

I might be occurred to M. Comte that a fact may be studied through the medium of memory, not at the very moment of our perceiving it, but the moment after: and this is really the mode in which our best knowledge of our intellectual acts is generally acquired. We reflect on what we have been doing, when the act is past, but when its impression in the memory is still fresh.” (*Auguste Comte and Positivism*, apud James 1890, p. 188)

O endosso de James a esta teoria do conhecimento introspectivo como retrospecção está diretamente ligado à sua concepção dos processos internos do pensamento, tal como exposta no início do capítulo *The Stream of Thought*, onde as três primeiras características que James postula são:

- 1) Todo pensamento tende a ser parte de uma consciência pessoal;
- 2) Internamente a cada consciência pessoal, o pensamento está sempre mudando;
- 3) Internamente a cada consciência pessoal, o pensamento é sensivelmente contínuo.

A concepção dos processos de pensamento como mudanças contínuas está exemplificado no exame dos enunciados de atribuição de estados mentais feitos em primeira pessoa, no tempo presente:

"There are, it is true, cases which we appear to be naming our present feeling, and so to be experiencing and observing the same inner fact at a single stroke, as when we say "I feel tired", "I am angry", etc. But these are illusory, and a little attention unmask the illusion. The present conscious state, when I say "I feel tired", is not the direct state of tire; when I say "I feel angry", it is not the direct state of anger. It is the state of *saying-I-feel-tired*, of *saying-I-feel-angry*,- entirely different matters, so different that the fatigue and anger apparently included in them are considerable modifications of the fatigue and anger directly felt the previous instant.(James 1890, p. 188-189, sublinhado meu)

Ao argumentar deste modo, James está apenas extraindo consequências e, portanto, ampliando, o que chamei acima da tese da não-composicionalidade, defendida por Brentano.

Uma das intuições básicas por trás da concepção da introspecção como retrospecção é este caráter episódico, de mudanças contínuas dos processos de pensamento, seu caráter caleidoscópico. E esta concepção pode ser encontrada

em vários momentos da psicologia e da filosofia posterior a James, como ilustram as citações abaixo de G. Daves Hicks, publicada em 1927, e da influente obra de Ryle, *The Concept of Mind*, publicada em 1949:

It is probably true that, in the vast majority of cases, a reflective consideration of a state of mind, a psychological scrutiny of it, is only possible when the original experience itself is past, and that, consequently, introspection is usually, at least, retrospection. (G. Daves Hicks, *Proceedings of the Aristotelian Society, Supplementary Volumes, Vol. 7, Mind, Objectivity and Fact* (1927), pp. 55-97)

Part, then, of what people have in mind, when they speak familiarly of introspecting, is this authentic process of retrospection. But there is nothing intrinsically ghostly about the objects of retrospection. In the same way that I can catch myself daydreaming, I can catch myself scratching; in the same way that I can catch myself engaged in a piece of silent soliloquy, I can catch myself saying something aloud. (Ryle 1949, p. 166)

- *The Experimental Method* – Apesar do fato de James ter sido o responsável por um dos primeiros laboratórios de psicologia experimental nos Estados Unidos, ele apresenta com certo desdém neste subcapítulo os métodos experimentais desenvolvidos na Alemanha, em especial por Wundt. Os pontos que merecem destaque são: a quantificação estatística e o controle laboratorial das condições sob as quais eram praticados os experimentos de “percepções internas”, como Wundt as denominava, usando a distinção de Brentano (Brentano 1874) entre “percepção interna” (*Wahrnehmung*) e “observação interna” (*Beobachtung*).

- *The Comparative Method* – Quando as características descritas pela psicologia da introspecção normal são colocadas em questão, resta ao pesquisador comparar estes resultados com casos desviantes analisados pela própria psicologia ou por outras áreas do conhecimento:

So it has come to pass that instincts of animals are ransacked to throw light on our own; and that reasoning faculties of bees and ants, the minds of savages, infants, madmen, idiots, the deaf and blind, criminals and eccentrics, are all invoked in support of this or that special theory about some part of our own mental life. The history of sciences, moral and political institutions, and languages, as types of mental product, are pressed into the same service. (James, 1890, p. 193)

Um ponto interessante aqui deve ser destacado. James, claramente, afirma que o método comparativo é um “suplemento” aos métodos introspectivo e experimental. Por razões óbvias, os estudos de psicologia anormal e psicologia animal não podem ser baseados em observações introspectivas. Ou seja, entre os métodos legítimos da pesquisa psicológica James deixa um espaço claramente demarcado para as metodologias behaviouristas ou funcionalistas que, mesmo ainda não delineadas quando James escrevia *The Principles*, estavam com seu lugar previsto.

Que traços da psicologia de James podem ser importantes para os temas da psicologia moral? James caracteriza *The Principles* como uma “coleção de ilustrações da dificuldade de descobrir por introspecção direta exatamente o que nossos sentimentos (*feelings*) e suas relações são” (James 1890, p. 191, meu sublinhado), e é preciso dar poucos passos analíticos para incluir as emoções no subconjunto de estados mentais que James nomeia “sentimentos”. Sendo assim, poderíamos reescrever a frase de James como asseverando das dificuldades de descobrir pela introspecção o que são nossas emoções. Se, além disso, lembrarmos da importância atribuída por neoemotivistas como Jesse Prinz à teoria das emoções de matiz não-cognitivista desenvolvida por James em seu artigo “*What is a Emotion?*”, publicado em *Mind*, em 1884, e desenvolvida no capítulo XXV de *The Principles* (posteriormente ficou conhecida como teoria James-Lange, em razão do trabalho de Carl. G. Lange - *Om Sindsbevaegelser: Et Psyko-Fysiologisk Studie* - publicado um ano depois, em 1895), então teríamos uma primeira pista da importância de elementos da visão de James para teorias

contemporâneas em psicologia moral. Prinz apresenta a teoria de James-Lange nos seguintes termos:

The claim that emotions lack intentionality is closely associated with noncognitive theories, but it is not essential to them. Consider the theory advanced by William James (1884) and Carl Lange (1885). They argue that emotions are feelings of patterned changes in the body. Lange emphasizes changes in blood vasculature. James is more inclusive. He mentions facial expressions, musculoskeletal changes, activity in visceral organs, and even stereotyped behaviors.

In *The Principles of Psychology* James quotes from the detailed observations of Darwin and others who took note of the intimate link and steady correlation between affect and bodily response – the pallor of fear, the clenched fists of anger, the throat lump of grief, and the toothy grin of joy. These responses, from flush to phlegm and fiber, are ordinarily thought to arise after an emotion is elicited. We say that sadness makes us cry, and not the reverse. According to James and Lange, folk psychology gets things backward. Emotions are internal states that register bodily changes. The central argument for both of them involves an exercise in mental subtraction. Imagine an intense emotion and then systematically eliminate every bodily feeling associated with that state. When the last bodily feeling is removed, there will be nothing left to the state that one would call an emotion. Emotions are felt perceptions of bodily changes.

(Prinz 2007, p.53)

Além da tese de que emoções não são objeto exclusivo de exame introspectivo, já que sua individuação depende de percepção de mudanças corporais, poderíamos acrescentar a este quadro outra tese de James, que nos parece associada a esta (apesar de James não apresentá-las conjuntamente), sobre o papel que a percepção que temos de nosso próprio corpo joga na construção do conceito de identidade pessoal. Este ponto foi destacado por Galen Strawson – talvez o maior entusiasta de James no contexto da filosofia contemporânea – quando ele alinha James, Nietzsche, Wundt, Dewey, Feuerbach, Fichte, Platner e Bradley no reconhecimento que a percepção que temos de nosso corpo – o que Bradley nomeava *coenaesthesia* (o que podemos traduzir como propriocepção, ou a totalidade das diferentes sensações que temos do funcionamento e posição das partes de nosso corpo) – é a origem da autoconsciência, mas mesmo assim, podemos ainda distinguir em nosso

conceito de identidade pessoal um papel desempenhado pelo sujeito (*self*) que é destacado, separado da pertença a um corpo:

For, whatever the thought we are criticising may think about its present self, that self comes to its acquaintance, or is actually felt, with warmth and intimacy. Of course this is the case with the *bodily* part of it; we feel the whole cubic mass of our body all the while, it gives us an unceasing sense of personal existence. Equally do we feel the inner 'nucleus of the spiritual self', either in the shape of yon physiological adjustments, or (adopting the universal psychological belief), in that of the pure activity of our thought taking place as such...The character of...warmth and intimacy...in the present self...reduces itself to either of two things – something in the feeling which we have of the thought itself, as thinking, or else the feeling of the body's actual existence at the moment, – or finally to both. (James 1890, p. 316, meu sublinhado)

Galen Strawson coloca esta distinção nos seguintes termos:

Our background awareness of our bodies is important, but this is wholly compatible with our regularly experiencing ourselves primarily or centrally as inner conscious presences who are not the same thing as human beings considered as a whole; and although background awareness of body is indeed experience of the body, this doesn't prevent it from feeding or grounding SELF-EXPERIENCE, a sense of self that presents the self primarily as a distinctively (and in some cases purely) mental entity. (Strawson 2009, p. 28)

Para Galen Strawson, o ponto importante nesta distinção é exatamente capturar o que está envolvido em ter a experiência do sujeito como uma “entidade distintivamente mental”. Para nós, no contexto da presente discussão, o importante é destacar a relação entre a teoria das emoções de James, a saber, emoções como sensações de alterações corporais, com a tese do que James chama a “parte corporal” da identidade pessoal. No entanto, por mais que este aspecto da psicologia de James seja útil para a psicologia moral do neoemotivismo, esta seria uma utilização posterior das teorias de James em um quadro moral que não seria propriamente adequado à complexidade das reflexões de James sobre moralidade e sobre a noção de Self. Esta dissertação,

certamente, deve defender a tese segundo a qual toda teoria moral apresenta alguma pressuposição sobre o funcionamento da atribuição e autoatribuição de estados mentais, mais especificamente, pressupõe teses sobre atribuição e autoatribuição de atitudes proposicionais. Agora, seria justo afirmar que a aceitação, a adoção consciente de uma teoria psicológica, seja ela qual for, teria consequências para que tipo de teoria moral deveria ser aceita?

William James não é considerado um pensador dedicado às questões da moralidade, não encontramos o nome de James entre os textos clássicos do pensamento ético, não associamos o nome de James a nenhuma teoria moral. No próprio prefácio do *The Principles*, nos é dito que “the important subjects of pleasure and pain, and moral and aesthetic feeling and judgements” foram deixados de lado. Um dos poucos textos onde James diretamente tratou de conceitos morais é *The Moral Philosopher and the Moral Life*, publicado pela primeira vez em 1897 em *Will to Believe and Other Essays in Popular Philosophy*. Sendo assim, seria no mínimo heterodoxo defender a obra de James como tendo alguma importância destacada para as questões próprias da filosofia moral, ou que James pudesse nos dizer algo de valor sobre a moralidade que outros pensadores já não tenham dito. Mas neste ponto vou me permitir uma modesta heterodoxia. E a razão disto é minha propensão a responder afirmativamente a última pergunta feita no parágrafo anterior. Sim, acredito que diferentes teorias sobre o funcionamento de conceitos psicológicos acarretam diferentes consequências para aceitação de teorias morais específicas. E aqui temos um ponto a favor: James não foi apenas “um verdadeiro ser humano”, mimetizando aqui os elogios de Wittgenstein; James nos deixou uma obra de grande vigor intelectual, em especial ele foi, na condição de um dos pais fundadores da psicologia, um observador especialmente atento da experiência humana.

Não seria por outra razão que Stanley Cavell elenca *The Varieties of Religious Experience* entre os textos exemplares do Perfeccionismo Emersoniano:

A text that I have left unmentioned and that bears distinctly on Emersonian perfectionism is William James's still marvelous, brave *Varieties of Religious Experience*, which challenges both philosophy and religion in its faithfulness to human experience. (Cavell 2004, p.17)

Não vou me aventurar aqui na reconstrução do que poderia ser uma teoria moral *à la* William James, esta tarefa estaria muito além das minhas capacidades. Me limitarei a indicar que qualquer reconstrução deste tipo precisaria dar conta da multiplicidade de pontos de vista que uma teoria moral deveria adotar para explicar de forma aceitável as ações morais, em um modelo minimamente compatível com teses de *The Principles*. Esta multiplicidade deveria espelhar os três aspectos do que James classifica como "the empirical life of Self". No longo capítulo *The Consciousness of the Self*, James resume sua explicação do que está em jogo no conceito de amor próprio (*self-love*) usando a seguinte tabela:

	MATERIAL	SOCIAL	SPIRITUAL
SELF-SEEKING	Bodily Appetites and Instincts Love of Adornment, Foppery, Acquisitiveness, Constructiveness Love of Home, etc.	Desire to please, be noticed, admired, etc. Sociability, Emulation, Envy, Love, Pursuit of Honor, Ambition, etc.	Intellectual, Moral and Religious Aspiration, Conscientiousness
SELF-ESTIMATION	Personal Vanity, Modesty, etc. Pride of Wealth, Fear of Poverty	Social and Family Pride, Vainglory, Snobbery, Humility, Shame, etc.	Sense of Moral or Mental Superiority, Purity, etc. Sense of Inferiority or of Guilt

(James 1890, p.313)

Os três aspectos a que me referi são as características (1) materiais, (2) sociais e (3) espirituais do sujeito (*self*) empírico jamesoniano. Em relação aos aspectos materiais, James estaria de acordo em aceitar apenas como parte de uma teoria moral compatível com sua psicologia, explicações de cunho utilitarista empregando os conceitos de dor e prazer, relacionando-os aos "apetites corporais e instintos". Em *The Moral Philosopher and the Moral Life*,

quando James formula uma resposta possível para o que ele denomina a “questão psicológica” sobre a “origem histórica de nossas ideias e juízos morais”, ele concorda com Bentham e Mill:

I will therefore only express dogmatically my own belief, which is this, — that the Benthams, the Mills, and the Barns have done a lasting service in taking so many of our human ideals and showing how they must have arisen from the association with acts of simple bodily pleasures and reliefs from pain. Association with many remote pleasures will unquestionably make a thing significant of goodness in our minds; and the more vaguely the goodness is conceived of, the more mysterious will its source appear to be. (James 1897)

Uma explicação moral utilitarista, empregando os conceitos de prazer e dor, estaria vinculada ao que James chama “bodily self-love”, constituindo um dos elementos do conceito de amor próprio.

Mas nossa imagem da moralidade jamesoniana ainda estaria incompleta se ficássemos apenas restritos ao “bodily self-love”. Temos um interesse manifesto pela opinião que outras pessoas têm a nosso respeito. É somente na relação com os outros que os conceitos de inveja, vergonha e orgulho – apenas para ficarmos com alguns dos conceitos centrais de qualquer explicação moral – apresentam seu significado. Em um mundo de “solidão moral”, onde um único sujeito habitasse o universo “contendo apenas fatos físicos e químicos, e existindo desde a eternidade sem um Deus” (James 1897), como apresentado no experimento mental de James em *The Moral Philosopher and the Moral Life*, estes conceitos morais não existiriam:

Neither moral relations nor the moral law can swing *in vacuo*. Their only habitat can be a mind which feels them; and no world composed of merely physical facts can possibly be a world to which ethical propositions apply. (James 1897)

Esta esfera intersubjetiva constitui o espaço próprio ao “self social”, ou do que James chama “meu Eu histórico” (James 1890, p. 306), onde nossas relações com outras pessoas são definidas na convivência em sociedade ou no grupo familiar.

O último aspecto que uma visão jamesoniana da moral deve contemplar diz respeito ao “self espiritual”, mais diretamente ao que James chama de “aspirações morais e religiosas”, “sentido de superioridade moral” ou “sentido de inferioridade e culpa”. Neste ponto, gostaria de sugerir que o conceito de “alma enferma” (*sick soul*) que James desenvolveu em *The Varieties of Religious Experience*, não captura apenas um traço recorrente da fenomenologia de uma experiência religiosa, mas define um tipo de personagem com forte matiz moral.

A “alma enferma” experimenta o mundo trespassado pela onipresença do mal, estamos permanentemente em estado de pecado no mundo, na iminência do erro e sofrendo na melancolia de nosso desajuste com a ordem natural. A perspectiva da “alma enferma” descreve a situação daquelas pessoas “que não conseguem livrar-se tão rapidamente do fardo da consciência do mal, mas estão congenitamente fadados a sofrer pela sua presença”. (James 1902/2002, p. 108). Este sofrimento manifesta-se em uma forma de melancolia religiosa que James descreve como “perda passiva de todo apetite pelos valores da vida” (James 1902/2002, p. 120). Para o melancólico, “o mundo parece agora distante, estranho, sinistro, bizarro. Sua cor se foi, seu hálito é frio, não há o brilho da curiosidade em seus olhos.” (James 1902/2002, p. 121).

O próprio James reconheceu em *The Varieties* que não discutiria por palavras com aqueles que quisessem reconhecer em algumas de suas discussões, temas propriamente morais:

But if you say this, it will only show the more plainly how much the question of definition tends to become a dispute about names. Rather than prolong such a dispute, I am willing to accept almost any name for the personal religion of which I propose to treat. Call it conscience or morality, if you yourselves prefer, and not religion – under either name it will be equally worthy of our study. (James 1902/2012, p. 29)

Por mais que eu reconheça, com James, que alguns dos conceitos desenvolvidos em *The Varieties* (alguns exemplos dos conceitos que tenho em mente são: “*divided self*”, “*healthy-mindedness*” por oposição a “*sick soul*”,

“conversion” e “saintliness”) apresentam traços característicos de fenômenos exclusivos da experiência religiosa, traços que “a moralidade pura e simples não contém” (James 1902/2012, p. 29). Mesmo assim, quero reconhecer na importância que James atribui ao conceito de “alma enferma” uma dimensão moral, na medida em que o sujeito afetado por esta sensibilidade particular está imerso em uma situação de confronto absoluto, de crise emocional que terá sua resolução religiosa na forma de uma conversão, e cuja contrapartida moral seria uma espécie de resolução na adoção de determinados valores.

As ressonâncias morais do conceito de “alma enferma” me fazem lembrar da reivindicação de Cavell da pertença de *The Varieties* ao rol do textos de inspiração perfeccionista. A crise da melancolia religiosa aproxima-se da “quiet desperation” que Thoreau atribui às vidas de seus vizinhos. Aproxima-se também da “silent melancholy” emersoniana. Quando James analisa a figura de Leo Tolstoy, tal como transparece no texto autobiográfico de *Confissão*, publicado em 1882, lemos o seguinte:

As I interpret his melancholy, then, it was not merely an accidental vitiation of his humors, though it was doubtless also that. It was logically called for by the clash between his inner character and his outer activities and aims. Although a literary artist, Tolstoy was one of those primitive oaks of men to whom the superfluities and insincerities, the cupidities, complications, and cruelties of our polite civilization are profoundly unsatisfying, and for whom the eternal veracities lie with more natural and animal things. His crisis was the getting of his soul in order, the discovery of its genuine habitat and vocation, the escape from falsehoods into what for him were ways of truth. (James 1902/2012, página 147, meus sublinhados)

Esta caracterização de James da melancolia religiosa de Tolstoy apresenta várias características em comum com a apresentação que Cavell faz do perfeccionismo moral em *Cities of Words*:

This aspect or moment of morality – in which a crisis forces an examination of one’s life that calls for a transformation or reorienting of it – is the province of what I emphasize as moral perfectionism. I do not conceive of this as an alternative to Kantianism or utilitarianism (...)but rather as emphasizing that aspect of moral choice having to do, as it is sometimes put, with being true to oneself...(Cavell 2004, p. 11, meus sublinhados)

“Colocar a alma em ordem” poderia ser o mote perfeccionista por trás do conceito de “alma enferma”, representando um aspecto do “spiritual self”, na classificação tripartite do sujeito empírico jamesoniano.

2.1.3 James Sully

O psicólogo inglês James Sully (1842 - 1923), apesar de amigo relativamente íntimo de William James, não tinha uma opinião muito caridosa a respeito de *The Principles*. Além de queixas em relação ao estilo, não conseguia ver um plano no que lhe parecia apenas uma coleção de artigos. Talvez esta fosse apenas uma reação às poucas críticas que James dirige a ele em momentos do seu texto, mesmo que permeadas aos frequentes elogios. Mas Sully aparece em *The Principles* em um momento particularmente importante para os nossos propósitos em uma passagem parcialmente citada acima, onde James caracteriza a introspecção como falível:

Mr. Sully, in his work on *Illusions*, has a chapter on those of Introspection from which we might now quote. But, since the rest of this volume will be little more than a collection of illustrations of the difficulty of discovering by direct introspection exactly what our feelings and their relations are, we need not to anticipate our own future details, but just state our general conclusion that *introspection is difficult and fallible; and that difficulty is simply that of all observation of whatever kind.* (James 1890, p. 191)

De fato, James não cita *Illusions*, pelo menos não neste contexto de discussão dos *The Principles*. Todavia, ao revisar o que os psicólogos do final do século XIX pensaram sobre introspecção, me parece claramente que Sully foi o primeiro a adotar a posição que Schwitzgebel (2001) considera tão rara, a saber, dar peso metodológico para os processos introspectivos em uma caracterização adequada do mental, ao mesmo tempo em que enfatiza a falibilidade destes processos. E esta tarefa ele executou no capítulo chamado “*Illusions of Introspection*” de *Illusions*, editado pela primeira vez em 1881.

Considerando que estejamos corretos na afirmação deste pioneirismo de Sully, é esperada uma certa flutuação terminológica, típica dos cenários de discussão ainda não estabelecidos dentro da literatura. Não creio que Sully tenha os melhores exemplos para caracterizar o que ele chama de “*ilusões da introspecção*”, mas alguns de seus primeiros *insights* trazem elementos interessantes, em especial, naquilo que antecipam determinadas posturas que

James irá expressar em *The Principles*. Como a composição do texto de James tomou praticamente 12 anos de seu trabalho a partir de 1878, não seria descabido supor que esta posição básica, compartilhada por James e Sully sobre a relevância e falibilidade da introspecção fosse um terreno comum a estes dois autores, ou mesmo supor que pudesse ser estendido a uma comunidade mais ampla de psicólogos ingleses com os quais James mantinha contato intelectual e relações de amizade, mesmo antes da publicação de *The Principles*.

Dentre os pontos que merecem destaque neste texto de Sully, sua distinção entre ilusões da introspecção passivas e ativas é especialmente importante. Sully descreve uma fenomenologia caleidoscópica e fugidia dos estados mentais que, de certa forma, assemelha-se ao conceito jamesoniano de “stream of thought”. Ilusões passivas são aquelas causadas pela complexidade, transitoriedade e variedade de nossa experiência interna:

Under ordinary circumstances, external impressions persist so that they can be transfixed by a deliberate act of attention, and objects rarely flit over the external scene so rapidly as to allow us no time for a careful recognition of the impression. Not so in the case of the internal region of mind. The composite states of consciousness just described never remain perfectly uniform for the shortest conceivable duration. They change continually, just as the contents of the kaleidoscope vary with every shake of the instrument. Thus, one shade of feeling runs into another in such a way that it is often impossible to detect its exact quality; and even when the character of the feeling does not change, its intensity is undergoing alterations so that an accurate observation of its quantity is impracticable. Also, in this unstable shifting internal scene features may appear for a duration too short to allow of close recognition. In this way it happens that we cannot sharply divide the feeling of the moment from its antecedents and its consequents. If, now, we take these facts in connection with what has been said above respecting the nature of the process of introspection, the probability of error will be made sufficiently clear. (Sully 1881, p. 67)

Ilusões ativas seriam tributadas às ideias pré-concebidas que temos em relação ao que descobriríamos sobre nossos estados mentais se estivéssemos envolvidos em um processo de exame introspectivo. As variedades do

autoengano são os exemplos emblemáticos deste tipo de ilusão, como fica claro nesta alusão indireta ao bispo Joseph Butler:

The great region for the illustration of these active illusions is that of the moral and religious life. With respect to our real motives, our dominant aspirations, and our highest emotional experiences, we are greatly liable to deceive ourselves. The moralist and the theologian have clearly recognized the possibilities of self-deception in matters of feeling and impulse. To them it is no mystery that the human heart should mistake the fictitious for the real, the momentary and evanescent for the abiding. (Sully 1881, p.69)

A identificação que Sully faz entre seu conceito de “ilusão ativa” e “as possibilidades de autoengano em questões de sentimento e impulso” é especialmente interessante e nos dá a possibilidade de tratar, mesmo que lateralmente, de um tema fundamental: a discussão sobre autoengano em contextos morais. Quem seriam os moralistas e teólogos que Sully refere nesta passagem? Como indiquei acima, a primeira suspeita recai sobre Joseph Butler, que atende às duas caracterizações. Butler, que escreve em uma tradição de pensadores cristãos entre os séculos XVII e XVIII, deve ser lembrado em primeira mão por sua influência no pensamento moral subsequente, enquanto os demais autores desta tradição, nomes como Daniel Dyke, Richard Baxter e John Mason, não gozaram do mesmo prestígio. ¹ A influência de Butler é especialmente marcada sobre Adam Smith e Kant, não por acaso, dois autores que trataram do tema do autoengano em seus escritos morais. Me parece especialmente claro que a possibilidade do autoengano serve como evidência para a precariedade de nossa experiência introspectiva quando o que está em jogo é a descrição de nossos interesses, motivos, “sentimentos e impulsos” e

¹ Devo a indicação destes autores à leitura da dissertação de mestrado de José Eduardo Freitas Porcher – “Autoengano e delírio. Dois ensaios sobre crença e racionalidade” – UFRGS, 2011.

razões que levam-nos a agir em situações de conflito moral ou mesmo em decisões cotidianas.

O espaço de nossos motivos é, por vezes, suficientemente opaco para permitir que tenhamos dúvidas sobre o quão eficiente é o controle que exercemos sobre nossas decisões conscientes. Como estamos na companhia de teólogos, não é exagero dizer que o tema é bíblico, como atesta Jeremias 17/9: “Enganoso é o coração, mais do que todas as coisas, e perverso; quem o conhecerá?”. Agora, mesmo concedendo este ponto, por qual razão o tema do autoengano deveria ter relevo no contexto da psicologia moral? Uma resposta para esta questão será desenvolvida ao final da secção 4.1., quando identificarmos as características do que chamaremos de visão não-relacional do argumento moral.

2.1.4 David Armstrong

A formulação contemporânea das teorias do sentido interno conceitualmente mais próxima ao modelo lockiano foi apresentada em 1968 por David Armstrong em *A Materialist Theory of the Mind*, onde encontramos a seguinte definição de “introspecção”:

If we make the materialist identification of mental states with material states of the brain, we can say that introspection is a self-scanning process in the brain. (Armstrong 1968, p. 324)

O antecedente da expressão condicional afirma a tese fisicalista que está sendo defendida nesta obra segundo a qual não há razões lógicas ou filosóficas para negar que estados mentais são estados ou processos físicos do sistema nervoso central. Além da identificação de estados mentais com estados físicos do cérebro, Armstrong associa estados mentais com disposições para determinados comportamentos, de onde temos uma definição comparativa dos processos introspectivos envolvidos com a consciência:

If I perceive a physical situation, then we have an inner mental state ‘directed’ in a certain way towards a certain physical situation. (...) Now if I am aware, not only of the physical situation, but also of the fact that I am perceiving, then we have a further mental state ‘directed’ in the same sort of way towards the original mental state. And if this further mental state, which *qua* mental state is simply a state of the person apt for the production of certain behavior, can be contingently identified with a state of the brain, it will be a process in which one part of the brain scans another part of the brain. (...) Consciousness, or experience, then (...) is simply awareness of our own state of mind. The technical term for such awareness of our own mental state is ‘introspection’ or ‘introspective awareness’. (Armstrong 1968, p. 94-5)

No conseqüente da expressão condicional citada acima, temos a afirmação da introspecção como um processo físico que ocorre no cérebro. É tratada como uma questão controversa entre teóricos do sentido interno, como Armstrong, Lycan e Alvin Goldman, se existe algum órgão fisicamente

responsável pela introspecção ou se devemos construir os processos introspectivos através de um “idioma das faculdades”, ao estilo de Kant, por exemplo. Sobre este ponto, que reaparecerá no tratamento dos argumentos de Shoemaker contra as teorias do sentido interno (capítulo 3.1), Armstrong tem um argumento sensivelmente fraco para apoiar a analogia entre sentido externo e interno:

It is sometimes argued that introspection cannot be compared to sense-perception because no sense-organ is involved. We say ‘I see with my eyes’, but there is nothing with which I can say that I discover I am thinking. Now I do not believe that this objection would carry much weight, even if the difference from sense-perception could be made out. But in any case there is one sort of sense-perception where we do not say that we perceive with anything: bodily perception. When I become aware that I am hot, or that my limbs are moving, and I do not gain this knowledge by touch, there is no organ that I can say that I perceive these things with.(Armstrong 1968, p. 95-6, meu sublinhado)

O recurso a propriocepção (que é como entendo o que Armstrong está chamando aqui de “percepção corporal”) como um caso onde percebemos sem o auxílio de um órgão sensorial específico, mesmo excluindo o tato, como Armstrong o faz em seu argumento, é empiricamente falso. Pois mesmo que desconsideremos o tato, existem vários outros fatores fisiológicos envolvidos na propriocepção, como receptores nervosos nas fibras musculares ou estruturas sensoriais do ouvido interno, o que desautoriza a propriocepção ou percepção corporal como um modelo adequado para explicar os processos introspectivos, pois não é correto que possamos defini-la como uma percepção sem o concurso de um ou mais órgãos com funções fisiológicas particulares.

Outro ponto importante para a teoria de Armstrong é sua definição de “estado mental” como “um estado da pessoa apto para a produção de um certo comportamento”. Produzir um comportamento deve ser entendido aqui como uma relação causal. Retomemos a definição tal como Armstrong originalmente a formula: “The concept of a mental state is primarily the concept of *a state of the person apt for bringing about a certain sort of behavior*” (Armstrong 1968, p. 82). Por

esta definição, somada à tese de Armstrong de que causa e efeito possuem “existências distintas”, podemos pensar que o estado mental envolvido no ato introspectar e o estado mental introspectado mantêm entre si uma relação causal e, logo, contingente:

The further assumptions I will make about the nature of the causal relations in this work will only be two in number. In the first place, I will assume that the cause and its effect are ‘distinct existences’, so that the existence of the cause not logically imply the existence of the effect, or vice versa. In the second place, I will assume that if a sequence is a causal one, then it is a sequence that falls under some law. (Armstrong 1968, p. 83, meu sublinhado)

Além desta conclusão sobre a relação contingente entre estado de introspectar e estado introspectado, que será criticada por Shoemaker como veremos na seção 3.1.6, Armstrong é sensível a outra crítica que ele mesmo levanta e procura resolver. Vejamos os termos em que ele coloca a questão:

It may still be objected that if we accept the Causal analysis of mental states (...) introspection cannot possibly be compared to sense-perception, even bodily sense-perception. For consider what would be introspectively aware of. (i) We would be directly aware of an extraordinary abstract, and purely relational, state of affairs. We would be directly aware that something of whose non-relational properties we had no direct awareness at all was operating to produce certain behavior. (ii) The awareness would be a direct awareness of *causes*: a direct awareness that certain behaviour was being produced. (...) Now, it may be objected, this is incredibly far from the detailed awareness of intrinsic properties of objects that is yielded by sense-perception. (Armstrong 1968, p. 96)

A dificuldade apresentada neste trecho também foi examinada por Shoemaker e apresentaremos seu argumento em 3.1.4. A solução que o próprio Armstrong oferece aqui é interessante, mas discutível. Para resumir a dificuldade, podemos dizer que nos casos paradigmáticos de percepção somos conscientes de propriedades intrínsecas dos objetos, enquanto no caso da introspecção, de acordo com a definição de “estado mental” proposta por Armstrong, estaríamos conscientes de causas de comportamento próximos ou

ocorrentes, logo, estaríamos conscientes de estados de coisas “abstratos e puramente relacionais”. A solução de Armstrong é oferecer um caso de percepção sensorial, onde o que percebemos não é uma propriedade ou conjunto de propriedade intrínsecas de objetos, mas uma relação causal:

Suppose I feel a pressure in the small of my back. What am I aware of? It may be that I aware of no more than this: something I know not what is pressing upon my back. I might say it was something material, but what is a material object in this context except ‘that which is capable of exerting pressure’? I might not even know whether it was something solid, something liquid (such as a jet of water), or something gaseous (a jet of air). My awareness of the object is simply awareness of ‘something which has the relation to me of pressing on me’. Here is a perceptual parallel to the abstract and relational awareness that is being attributed to ‘inner sense’. (Armstrong 1968, p.97)

Nesta solução, parece que temos uma situação que, infelizmente, é encontrada ou recorrente na argumentação filosófica. Para dar credibilidade a sua tese, Armstrong empobrece a experiência que nos pede para imaginar. Se algo pressiona minhas costas, e estou restrito apenas às minhas sensações táteis (o que já é um empobrecimento, pois poderia descartar a possibilidade de ser um jato de água pressionando minhas costas, pois não ouço o som da água caindo), tenho que desconsiderar fatores como intermitência, temperatura ou rugosidade do objeto que sinto pressionando minhas costas, para dar plausibilidade ao argumento de que nesta experiência tátil não tenho consciência de propriedades intrínsecas que, em uma situação normal, estariam fazendo parte do cenário perceptual.

2.2 Olhando para fora: Gareth Evans

Poucas passagens na filosofia contemporânea foram tão influentes e suscitaram tanta discussão quanto o trecho de *The Varieties of Reference* que vamos agora examinar:

In fact, I shall concentrate upon the ways we have of knowing what we *believe* and what we *experience*, for I believe that if we get these right, we shall have a good model of self-knowledge (or introspection) to follow in other cases. In particular, I shall quite avoid the idea of this kind of self-knowledge as a form of perception – mysterious in being incapable of delivering inaccurate results.

Wittgenstein is reported to have said in an Oxford discussion:

If a man says to me, *looking at the sky*, ‘I think it is going to rain, therefore I exist’, I do not understand him.

The contribution is certainly gnomic; but I think Wittgenstein was trying to undermine the temptation to adopt a Cartesian position, by forcing us to look more closely at the nature of our knowledge of our own mental properties, and, in particular, by forcing us to abandon the idea that it always involves an *inward* glance at the states and doings of something to which only the person himself has access. The crucial point is the one I have italicized: in 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 be there a third world war?’ I get myself in a position to answer the question whether I believe that *p* by putting into operation whatever procedure I have for answering the question whether *p*.(Evans 1982, p. 225)

O primeiro ponto que merece destaque é o seguinte: Evans considera sua explicação de como autoatribuímos *crenças* e *experiências* como ilustrativa para uma teoria da introspecção em geral. Podemos pensar, portanto, que seu tratamento de autoatribuição de crenças e experiências perceptuais pode ser explicativo para autoatribuição de outras atitudes proposicionais como “desejo que *p*”, “espero que *p*” e “suponho que *p*”, e que também possa ser explicativo ou esclarecedor para o modo com autoatribuímos traços de caráter como

covardia, honestidade ou perseverança, por exemplo. Concordando com Gallois 1996, Nichols e Stich 2003, Bar-On 2004, Goldman 2006 e Carruthers 2011, a possibilidade desta extensão da teoria não parece conceitualmente defensável.

O tratamento que Evans deu para autoatribuição de experiências não foi tão decisivo no desenvolvimento da literatura posterior quanto seu tratamento da autoatribuição de crenças. Duas exceções são o recente artigo de Alex Byrne, *Knowing What I See*, publicado na coletânea *Introspection and Consciousness* (2012), que examina as teses de Evans sobre autoatribuição de experiência visual e a concordância com as teses de Evans sobre autoatribuição de experiências expressa por Peter Carruthers em *The Opacity of Mind*, apesar de sua crítica da tese da transparência quando aplicada à autoatribuição de atitudes proposicionais.

A tese de Evans, segundo a qual responder a uma pergunta sobre minha crença em *p* demanda os mesmos procedimentos que uso para responder sobre o valor de verdade de *p*, ficou conhecida como tese da transparência e já havia sido expressa, como lembra Richard Moran 2001, por Roy Edgley em *Reason in Theory and Practice*:

My own present thinking, in contrast to the thinking of others, is transparent in the sense that I cannot distinguish the question 'Do I think that P?' from a question in which there is no essential reference to myself or my belief, namely 'Is it the case that P?' This does not of course mean that the correct answers to these questions must be the same; only I cannot distinguish them, for in giving my answer to the question 'Do I think that P?' I also give my answer, more or less tentative, to the question 'Is it the case that P?' (Edgley 1969, p.90)

O segundo ponto ao qual gostaríamos de dar relevo é o seguinte: há um ganho quando nos apercebemos que, ao menos em alguns casos, como nas autoatribuições de crença e experiências perceptuais, não precisamos pressupor nenhum tipo de percepção interna, nenhuma espécie de olhar interior para descrever corretamente um estado mental como sendo *meu* estado mental. Este

é o ponto de minimizar a “tentação cartesiana”. Evans insiste sobre isso, pelo menos em mais dois momentos adiante no texto:

However, in an important respect these two ways of gaining self-knowledge are similar: namely that neither conforms to the description ‘looking within’. In the case we have just been describing, the subject’s concentration, as with self-ascription of belief, is on the outside world: how does he, or would he, judge it to be? (Evans 1982, p. 230, sublinhado meu).

It is of the utmost importance to appreciate that in order to understand the self-ascription of experience we need to postulate no special faculty of inner sense or internal self-scanning. (Evans 1982, p.230, sublinhado meu)

É justificada a pretensão expressa por Evans de que seu exame da transparência da autoatribuição de crenças e experiências perceptuais seja relevante para o exame da autoatribuição de outras atitudes proposicionais? Bar-On, em *Speaking My Mind* (2004), coloca esta pergunta nos seguintes termos:

Although Evans introduces the idea of transparency in connection with self-ascriptions of beliefs, I think his observation can be generalized to other intentional avowals. If asked whether I am hoping or wishing that *p*, whether I prefer *x* to *y*, whether I am angry at or afraid of *z*, and so on, my attention would be directed at *p*, *x* and *y*, *z*, etc. For example, to say how I feel about an upcoming holiday, I would consider whether the holiday is likely to be fun. Asked whether I find my neighbor annoying, I would ponder her actions and render a verdict. Considering whether I am scared of the dog, I will think of the dog, and whether it is scary. In general, in addressing questions about what I think, believe, want, prefer, feel, and so on, I concern myself not with me and my states, but rather with the world outside myself. (Bar-On 2004, p. 106-107)

Usando a fórmula de Evans, se alguém me perguntasse “Você deseja que chova amanhã?”, não me parece sensato dizer que eu responderia a esta pergunta da mesma maneira, e usando dos mesmos procedimentos que usaria para responder à questão “Vai chover amanhã?”. Também não parece ter a mesma descrição procedimental nossa resposta para as questões “Você teme andar nesta montanha russa?” e “É temeroso/perigoso andar nesta montanha

rusa?”. O que estas indicações mostram é que a atitude proposicional da crença guarda uma relação conceitual com a descrição dos fatos acreditados que “esperar que p ”, “temer que p ” e “desejar que p ” não guardam com a descrição dos fatos esperados, temidos ou desejados. O ponto poderia ser melhor descrito dizendo que “Temo que p , mas não- p ”, “Desejo que p , mas não- p ”, “Espero que p , mas não- p ” não são construções que geram o mesmo tipo de absurdo que as construções que apresentam o paradoxo de Moore, seja na forma comissiva ou omissiva.

Vejamos o argumento de Peter Carruthers em *The Opacity of Mind* (2011), contra a extensão da tese da transparência para outras atitudes proposicionais:

One problem with this suggestion is that it appears to have only a limited range of application. For even if the proposal works for the case of judgment and belief, it is very hard to see how one might extend it to account for our knowledge of our own goals, decisions and intentions – let alone our knowledge of our own attitudes of wondering, supposing, fearing and so on. For in such cases it is doubtful whether there is any question about the world, finding the answer to which will tell me what I want, decide, intend, or fear. In the case of desire I can ask whether something is *good*, for example. But then the inference from, ‘X is good’ to, ‘I want X’ appears shaky. Likewise in the case of fear I can ask whether something is dangerous, or fearful. But again the inference from ‘X is fearful’ to, ‘I fear X’ seems equally unsound. I can perfectly well recognize that something is fearful without be afraid of it. (Carruthers 2011, p. 81)

Quando Carruthers afirma que a conclusão “eu quero X” a partir da premissa “X é bom” é instável, podemos generalizar o ponto, dizendo que não há uma relação necessária entre as duas proposições. Pode muito bem ser o caso que doar sangue seja uma boa atitude e, no entanto, que eu possa defender a posição de não querer doar sangue, ou que caminhar à noite por cemitérios seja uma experiência amedrontadora, mas que eu não tenha medo destas aventuras góticas.

Por outra razão, a saber, deixar espaço para o caráter expressivo das autoatribuições de estados mentais, Bar-On também traça limites ao método da transparência, oferecendo novos exemplos:

But now consider spontaneously volunteered avowals such as “I’d really like a cup of tea right now”, or “I am so worried about my dad” (declared unprompted), or “I am wondering what time it is” (said out of the blue). In this kind of case, no question is posed to me (by myself or by anyone else), which specifies a subject matter or object for my ‘transparent’ consideration. I simply pronounce on my current state of mind. While I agree that it is not fruitful to portray such avowals as the upshot of an ‘inward gaze’, I think it is equally implausible to suppose that they involve a look at the world outside us. Indeed, it is very unclear *where* in the outside world one would look to find out what one’s state of mind is in this type of case. (Bar-On 2004, p. 114-115)

Suponhamos que me fizessem as seguintes perguntas: “Você acredita que você é generoso/honesto/covarde?”. Neste caso, parece claro que não responderei a estas questões respondendo, primeiro, as perguntas “Sou generoso/honesto/covarde?”. Uma certa autoimagem que faço servirá para responder a primeira questão sem a necessidade de considerar a segunda. Para um sujeito suficientemente cínico, por exemplo, é possível imaginar uma situação onde ele responda negativamente a segunda questão, digamos, sobre sua honestidade, e afirmativamente a primeira questão. Por outro lado, podemos imaginar um sujeito de tal forma convicto de sua generosidade que consideraria uma afronta elencar os supostos fatos que justificariam a autoatribuição de generosidade. Nestes casos de predicados que descrevem traços de caráter, as situações que envolvem autoengano ou mesmo confabulação mostram evidência contrária ao procedimento da transparência sugerido por Evans.

Apesar das limitações da tese da transparência como explicação das autoatribuições de atitudes proposicionais ou estados mentais em geral, é importante guardarmos um ponto que será fundamental para as teorias racionalistas da introspecção, em particular para Richard Moran, a saber, a relação necessária entre o valor de verdade da minha autoatribuição de crença e a questão sobre o valor de verdade do conteúdo proposicional de minha crença:

With respect to belief, the claim of transparency is that from within the first-person perspective, I treat the question of my belief about P as equivalent to the question of the truth of P. What I think we can see now is that the basis for this equivalence hinges on

the role of deliberative considerations about one's attitudes. For what the 'logical' claim of transparency requires is the deferral of the theoretical question 'What do I believe' to the deliberative question 'What am I to believe?' And in the case of the attitude of belief, answering a deliberative question is a matter of determining what is true. (Moran 2001, p. 62-3)

Capítulo 3 – Modelo não-perceptual da Introspecção

“Our talk of introspection is metaphorical. I may see that another sees something, but not that I do; hear what he is listening to, but not perceive that I am hearing something. I can no more *look into* my mind than I can look into the mind of another. There is such a thing as introspection, but it is not a form of inner perception. Rather, it is a form of self-reflection in which one engages when trying to determine, for example, the nature of one's feelings (e.g. whether one loves someone); it is 'the calling up of memories; of imagined possible situations, and of the feelings that one would have if...' (PI §587). Such soul searching requires imagination and judgment, but no 'inner eye', for there is nothing to *perceive* - only to reflect on.” (Hacker 1997,p. 25)

Brie Gertler, em *Self-knowledge*, apresenta uma distinção entre teorias racionalistas e empiristas do autoconhecimento nos seguintes termos:

Empiricist theories claim that, in paradigm examples of self-knowledge, self-attributions are justified through a process in which one observes or detects one's own mental states. (...) Rationalist theories regard the capacity for self-knowledge of a particular type, critical self-knowledge of one's own attitudes, as a precondition for rational thought. Critical self-knowledge is essentially normative, in that it involves an exercise of one's epistemic *rights*. (...) Rationalism is sharply distinguished from empiricism by the claim that a principal type of self-knowledge is epistemically based in the normative dimensions of rational agency.(Gertler 2011, p. 255)

Grosso modo, os autores descritos no subcapítulo “Olhando para dentro” seriam classificados como empiristas pelo critério exposto por Gertler. Todos eles compartilham da analogia lockiana básica entre percepção externa e sentido interno. O espaço para abordagens racionalistas do autoconhecimento aparece neste capítulo “Modelo não-perceptual da Introspecção”, onde os principais autores apresentados serão Sydney Shoemaker e Richard Moran. Estes autores nos interessam por razões distintas. Shoemaker é o principal

crítico contemporâneo das teorias do sentido interno da introspecção. Fred Dretske, em artigo publicado recentemente (*Awareness and Authority: Skeptical Doubts about Self-Knowledge* em Smithies 2012), descreve as Royce Lectures *Self-knowledge and Inner Sense* (publicadas em Shoemaker 1996) como o caixão definitivo dos modelos perceptuais da introspecção.

Apresentar e comentar os argumentos de Shoemaker será um dos objetivos deste capítulo e o ponto central desta dissertação. Como afirmei na introdução, considero os resultados de Shoemaker como uma das conquistas definitivas da filosofia da mente contemporânea.

É reconhecido que o enfoque racionalista é especialmente desenhado para dar conta do conhecimento que supostamente temos de apenas um subconjunto de nossos estados mentais, em particular daqueles estados que apresentam um caráter ativo. Os processos envolvidos na decisão por cursos alternativos de ação, na descrição de nossas intenções de longo prazo e na autoatribuição de atitudes proposicionais ou traços de caráter são bons exemplos. A premissa básica é que temos uma autoridade em relação a estes casos pelo fato de que nos *comprometemos* em relação a eles. Posições que exploram este ponto podem ser encontradas em *Mind and Supermind* de Keith Frankish (2004), ou em *Self-Knowledge and Resentment* de Akeel Bilgrami (2006).

A distinção entre o caráter ativo ou passivo de nossos estados mentais precisa ser um pouco melhor desenvolvida. Retomemos a classificação dos estados ou processos mentais proposta por Gallois (1996) e utilizada no primeiro capítulo desta dissertação:

“(a) estados de ânimo, tais como depressão, euforia, ansiedade e desespero;

(b) traços de caráter tais como ser preguiçoso, generoso, alegre, corrupto, cruel, insensível, gentil, tolerante ou perspicaz;

(c) emoções como amor, aversão, raiva e medo;

(d) atitudes proposicionais como acreditar, temer, intencionar, desejar, esperar, saber, arrepende-se, perceber, compreender;

(e) sensações e experiências (visual, auditiva, tátil etc).”

Tomando esta classificação, os itens do grupo (e) são certamente estados passivos, no sentido tradicionalmente aceito de serem processos que nos acontecem. Em certo sentido, podemos argumentar na mesma linha para os itens do grupo (c). Por outro lado, os itens dos grupos (a), (b) e (d) são estados ou processos ativos, no sentido em que podemos racionalmente valorar e revisar o conteúdo destes estados ou o resultado destes processos.

A mais influente das publicações que adotou pontos de vista compatíveis com o que Gertler classifica como teorias racionalistas do autoconhecimento ou da introspecção foi *Authority and Estrangement* de Richard Moran (2001). Um breve comentário de alguns pontos da obra de Moran e a relação de dependência entre sua visão e posições de Shoemaker vai nos permitir introduzir uma das questões de fundo que esta dissertação procurará responder: qual a relação entre teorias da introspecção e alguns de nossos conceitos morais ordinários, como responsabilidade, culpa e ação intencional, por exemplo? Diferentes expressões para esta pergunta serão os temas do capítulo “Retratos da interioridade e conceitos morais”.

A principal tese defendida em *Authority and Estrangement* (2001) de Richard Moran é que o modelo perceptual do autoconhecimento (o que nesta dissertação estamos descrevendo como modelos perceptuais da introspecção) não capta traços essenciais da assimetria entre perspectivas de primeira e terceira pessoa. Quando a discussão epistemológica traça a peculiaridade da introspecção como uma questão de acesso (privilegiado ou não) ao espaço da vida mental consciente do sujeito, corre-se o risco de subdescrever o que há de propriamente relevante nesse processo:

The “internal theater” of Descartes (and Locke and Hume) and the long legacy of treating self-consciousness as a kind of inner perception is probably the most graphic expression of this approach, but the general tendency is broader than this. Nothing especially first-personal is captured by transferring the situation of a spectator from the outside to the inside, nor by construing the person as having any kind of especially good theoretical access to his own mind. (Moran 2001, p. 3, meu sublinhado)

Nesta dissertação estamos defendendo (seção 4.1) que os diferentes modelos das teorias perceptuais da introspecção dão suporte à uma visão não-relacional do argumento moral, e que esta visão deve ser deixada de lado em uma análise adequada desse argumento. A razão para isso é que esta visão não-relacional cai ou permanece em pé junto com as teorias perceptuais. Se tivermos boas razões para abandonar a analogia lockiana entre sentido interno e externo, então teremos as mesmas boas razões para desconsiderar a visão não-relacional.

Moran apresenta uma solução, a meu ver, mais conciliadora. A imagem do autoconhecimento gerada pelo modelo perceptual – o que Moran nomeia de “autoconhecimento como descoberta” – opõe-se a uma imagem racionalista do “autoconhecimento como resolução”, onde o que é destacado é o papel da agência racional na constituição dos estados mentais do sujeito. Moran não pretende que estes reflexos da imagem da introspecção como um “olho interior” devam ser totalmente deixados de lado, em prol de uma explicação alternativa para autoatribuição de atitudes proposicionais. As posturas da “descoberta” e da “resolução” são diferentes pontos de vista que o sujeito pode adotar e ambos são importantes para uma imagem do argumento moral. Determinados fenômenos que apresentam desafios para a psicologia moral, como autoengano e *akrasia*, podem ser vistos, inclusive, como o resultado do choque gerado pela adoção simultânea destas diferentes perspectivas sobre nós mesmos.

Tal como Shoemaker, Moran pretende oferecer uma explicação para os fenômenos do autoconhecimento que não recaia no que ambos caracterizam como uma imagem cartesiana² do mental:

...the fate of Cartesianism has been taken to be decisive for the very notion of first-person awareness, and recent philosophical work has been very creative in developing ways of describing the surface phenomena of first-person discourse that are deliberately deflationary of the claims of that discourse to be reporting any kind of genuine awareness. By contrast, I wish to defend a view of first-person awareness that seems it as both substantial, representing a genuine cognitive achievement, but which nonetheless breaks decisively with the Cartesian and empiricist legacy. As subsequent chapters will show, this entails not only rejection of the “inner eye” as applied to the mechanism of introspection, but an account of the general distortions of the purely theoretical or spectator’s stance toward the self (...). Being the person whose mental life is brought to self-consciousness involves a stance of agency beyond that of being a kind of expert witness. Thus the discussion taken up here moves from the epistemology of introspection to a set of issues in the moral psychology of the first-person. (Moran 2001, p.4, meu sublinhado)

Descrever o que está envolvido em considerar um determinado estado mental como *meu*, envolve aceitar que tenho uma responsabilidade por isso. É esta dimensão normativa que o racionalismo de Moran pretende capturar. Este enfoque na agência racional por certo não dá conta da diversidade de nossos estados mentais conscientes. É no mínimo estranho dizer que sou responsável por sentir um comichão no meu pé esquerdo ou pela sensação de ouvir os barulhos do tráfego. Tanto Moran como Tyler Burge (outro fundamental teórico de matiz racionalista) aceitam uma teoria híbrida do autoconhecimento:

I regard knowledge of one’s sensation as requiring separate treatment from of knowledge of one’s thoughts and attitudes. (Burge 1996, p.107)

² Ambos autores usam a expressão “imagem cartesiana” sem pretensão de que as características que atribuem a esta visão sejam facilmente identificáveis nos textos de Descartes. Nesta dissertação, tenho evitado este rótulo por entender que a epistemologia da introspecção lockiana é muito mais clara para os propósitos de identificar um modelo perceptual do que os textos de Descartes.

I will have comparatively little to say here about the case of sensations, which I believe raises issues for self-knowledge quite different from the case of attitudes of various kinds. (Moran 2001, p.9-10)

Para nossos fins, que o racionalismo tenha que aceitar explicações alternativas para consciência de nossas sensações não apresenta problema. A autoatribuição de atitudes proposicionais é o fenômeno que nos interessa, pois é neste processo que vislumbramos a intersecção entre teorias da introspecção e temas morais. Mas de que maneira a discussão sobre responsabilidade aparece neste cenário?

The special features of first-person awareness cannot be understood by thinking of it in terms of epistemic access (whether quasi-perceptual or not) to a special realm to which only one person has entry. Rather, we must think of it in terms of the special responsibilities the person has in virtue of the mental life in question being *his own*.(...)

It is modeling self-consciousness on the theoretical awareness of objects that obscures the specifically first-person character of the phenomenon, whether or not this theoretical perspective takes the specific form of the perceptual model of introspection. (Moran 2001, p. 32, meu sublinhado)

Vejamos agora, através da análise detalhada das *Royce Lectures* de Shoemaker, como o tema da agência racional está diretamente relacionado com a explicação do funcionamento da introspecção.

3.1 Sydney Shoemaker

Em filosofia, quando sabemos qual caminho não devemos tomar, já podemos considerar que temos um ganho significativo, mesmo que não saibamos ainda como fazer para chegar ao final. O saldo positivo das análises de Shoemaker sobre o autoconhecimento me parece ser este: a interdição da comparação entre introspecção e percepção sensorial. Mesmo que ele não proponha uma teoria substantiva que se apresente como uma explicação alternativa sobre o funcionamento dos processos introspectivos dentro de seu marco funcionalista, materialista e não-dualista, parece inegável que ele tenha desencorajado de maneira definitiva as teorias baseadas na analogia de inspiração lockiana entre sentido externo e interno. Os argumentos para isso estão concentrados nas conferências de “*Self-knowledge and ‘Inner Sense’*”, que constituem as *Royce Lectures*, publicadas em 1994, em especial nas conferências I e II (*The object perception model* e *The broad perceptual model*). Em parte, os argumentos utilizados nestes textos aparecem em artigos anteriores, especialmente em *Introspection and the Self* de 1986, *On Knowing One’s Own Mind* de 1988 e *First Person Access* de 1990.

A estratégia argumentativa das *Royce Lectures* é multi-dimensionada e de difícil resumo em poucas linhas. Começaremos pela análise que Shoemaker faz dos estereótipos da percepção sensorial que estão na base do que ele denomina modelo objeto-percepção (*object perception model*) e modelo perceptual amplo (*broad perceptual model*). A adequação das características dos processos introspectivos a estes modelos é que decidirá a questão sobre a pertinência das teorias perceptuais da introspecção, em especial sobre a pertinência de teorias do sentido interno tal como defendidas, por exemplo, por David Armstrong em *A Materialist Theory of the Mind* e William Lycan em *Consciousness and Experience*. Ou seja, o argumento leva a sério a analogia lockiana entre sentido externo e interno, estendendo-a ao limite.

Esta analogia não é exclusiva das teorias paradigmáticas do sentido interno, pois podemos encontrar traços desta analogia em boa parte das caracterizações

da introspecção na filosofia de inspiração cartesiana ou mesmo empirista, bem como no desenvolvimento da psicologia no começo do século XIX, como atesta a análise dos textos de Franz Brentano, James Sully e William James. Sendo assim, os argumentos de Shoemaker têm como alvo uma tradição extensa e, poderíamos dizer, hegemônica. Para fazer ruir esta analogia, Shoemaker lista oito características que constituem o estereótipo da percepção sensorial e as compara com os processos introspectivos (todas as citações traduzidas abaixo, que enunciam cada uma destas características, são de Shoemaker 1996, p. 204-206):

(1) Condição corpóreo-voluntária:

“Percepção sensorial requer a operação de um órgão da percepção cuja disposição está em alguma medida sob o controle voluntário do sujeito.”

(2) Condição experiencial:

“Percepção sensorial requer a ocorrência de experiências sensoriais, ou impressões sensoriais, que são distintas do objeto da percepção, e também distintas da crença perceptual (existindo alguma) que é formada.”

(3) Condição fato-objeto:

“Quando a percepção sensorial nos fornece a consciência de fatos, isto é, consciência *de que* tal e tal é o caso, ela o faz através da consciência de objetos.”

(4) Condição de identificação:

“Percepção sensorial proporciona ‘informação identificadora’ sobre o objeto da percepção. Quando alguém percebe, está apto a separar um objeto de outros, distinguindo-o dos outros pela informação, fornecida pela percepção, sobre suas propriedades tanto relacionais quanto não-relacionais.”

(5) Condição não-relacional:

“A percepção de objetos usualmente requer a percepção de suas propriedades intrínsecas, não-relacionais. Podemos perceber relações entre coisas que percebemos; mas não perceberíamos estas coisas em absoluto, e portanto não poderíamos perceber as relações entre elas, se elas não se apresentassem elas mesmas como tendo propriedades intrínsecas, não-relacionais.”

(6) Condição de atenção:

“Objetos da percepção são objetos potenciais da atenção.”

(7) Condição Causal:

“Crenças perceptuais são causalmente produzidas pelos objetos ou estados de coisas percebidos, através de um mecanismo causal que normalmente produz crenças que são verdadeiras.”

(8) Condição de independência:

“Os objetos e estados de coisas que são percebidos e sobre os quais a percepção fornece conhecimento, existem independentemente de serem percebidos, e, com certas exceções, independentemente de existirem coisas com a capacidade para percebê-los ou serem deles conscientes.”

As características de (1) a (8) constituem o *modelo objeto-percepção*, com ênfase especial sobre (3) a (6). As características (7) e (8) constituem, tomadas em separado, o *modelo perceptual amplo*. Vejamos, ponto a ponto, quais os argumentos de Shoemaker para mostrar que a comparação entre os processos introspectivos e as características (1) - (8) é fundada em uma analogia problemática.

3.1.1 Modelo objeto-percepção e condição corpóreo-voluntária

A primeira característica do estereótipo da percepção sensorial é a existência de um órgão (ou conjunto de órgãos) corporal que possibilita esta percepção. Para cada um dos cinco sentidos que o senso comum distingue, temos partes do corpo humano que cumprem esta função. Retina, nervos óticos, pavilhão auditivo, papilas gustativas, narinas e tecidos que constituem a pele, funcionando em conjunto com atividades cerebrais específicas, possibilitam que eu veja, sinta determinados aromas ou gostos e assim por diante. Posso fechar os olhos caso não queira ver algo, tapar as narinas para não sentir odores desagradáveis, ou mesmo usar fones de ouvido para amenizar os ruídos do tráfego.

Ter ou não experiências sensoriais particulares é algo que, em alguma medida, está sob meu poder, depende de uma decisão voluntária. Em relação a esta primeira característica a analogia entre processos perceptivos sensoriais e a introspecção não funciona. É ponto pacífico que não há nenhum órgão corporal responsável pelos processos introspectivos. Além disso, se comprarmos a metáfora visual da introspecção como uma espécie de “olho interior” teríamos que considerar que este olho não tem pálpebras. Para a introspecção, entendida de acordo com esta metáfora recorrente, não há nada análogo à atitude voluntária de fechar os olhos para não ver os conteúdos de nossa vida mental. O mais próximo disso seria o autoengano, mas não creio que seja factível caracterizá-lo como uma dimensão quase-sensorial e dependente da vontade do sujeito.

3.1.2 Condição experiencial

Quando percebo que a água do banho está muito quente, esta sensação me aparece de uma maneira particular, com uma fenomenologia específica. Posso descrevê-la como um ardor intenso e contínuo ou como um leve incômodo passageiro que se dissipa na medida em que permaneço mais tempo em contato com a água. Para cada experiência particular posso separar esta experiência do objeto que é experienciado (a água), das características deste objeto (sua temperatura) e das crenças causadas por esta experiência (“A água está muito quente” / “Consgo me acostumar com esta temperatura de banho”).

Para esta segunda característica do estereótipo perceptual, que chamamos aqui de condição experiencial, a analogia com processos introspectivos também falha. Não há para a introspecção uma fenomenologia particular. Uma “introspecção alegre” ou uma “introspecção generosa” seriam introspecções de estados de alegria ou de uma sensação de generosidade de parte do sujeito que tem estes estados e não uma fenomenologia específica *do ato de introspectar*. Sobre esta característica do caráter fenomenal da experiência sensorial, Shoemaker tem um ponto interessante. Eu posso olhar para um quadro de diferentes ângulos, posso experimentar a temperatura da água do banho primeiro nas costas, depois sobre a cabeça, posso me aproximar do forno para sentir mais de perto o cheiro do bolo de chocolate que senti lá da sala. Em resumo, posso modular minha experiência sensorial adotando diferentes perspectivas dentro do campo sensorial. Da introspecção, ao contrário, podemos dizer que é unidirecional. Não posso mudar a posição a partir da qual “observo” um estado ou processo mental. Mais do que isso, falar de “posição” neste caso é uma expressão enganadora. Não podemos falar de um lugar a partir do qual introspecto, no mesmo sentido que falamos de um lugar a partir do qual ouço a execução de uma sonata ou assisto uma encenação teatral.

A tese segundo a qual os processos introspectivos não possuem uma fenomenologia específica não é uma novidade no cenário da discussão filosófica sobre a consciência. William Lyons descreve o ponto nos seguintes termos:

If to introspect were in any literal sense to employ an inner sense, then it would have its own phenomenology. Just as the phenomenal experience of tasting is different from the phenomenal experience of touching and, linguistically, this difference provides different sets of unique predicates, so we should expect that if to introspect were to employ an inner sense, then introspecting would involve a sui generis experience with its own phenomenal qualities and would generate its own unique set of predicates when we came to describe these qualities in words. But when we come to describe our alleged introspectings, we do so in terms of what we introspect, not in terms of what it is like to introspect. Any experiential qualities in introspecting seem to be borrowed from the content of introspection. (Lyons 1986, p. 96, grifo meu)

3.1.3 Condição fato-objeto e condição de identificação

Para o exame das condições (3) e (4) Shoemaker distingue a introspecção do próprio sujeito (*self*) enquanto objeto do sentido interno (para usar uma dicção kantiana), da introspecção de seus estados: “Os candidatos para serem objetos (não-factuais) de percepção interna são, antes de tudo, o sujeito (*self*) ele próprio, aquele que é chamado ‘eu’, e, então, entidades mentais de vários tipos – sensações, sentimentos, pensamentos, crenças, desejos e assim por diante.” (Shoemaker 1996, 206). Os argumentos de Shoemaker para mostrar a inadequação das características de (3) e (4) dizem respeito à percepção interna que o próprio sujeito tem de si mesmo.

A inadequação do que chamamos condição fato-objeto está estreitamente ligada com a inexistência de um órgão corporal que seja responsável pela introspecção. Se quero formular alguma crença perceptual sobre o gosto do bolo de chocolate que está na geladeira do meu vizinho, deve ser possível que eu mantenha uma relação gustativa com este objeto. Devo, por exemplo, me dirigir até a casa do meu vizinho, pedir a ele uma fatia do bolo, colocá-lo em minha boca da maneira apropriada (não misturar a prova do bolo com a ingestão simultânea de uma porção de pimenta, por exemplo). Chamemos esta relação de *relação-de-provar-o-gosto*. Se não estabeleço esta relação com o bolo, não poderei formar nenhuma crença perceptual sobre o sabor do bolo.

Posso ter crenças sobre o sabor do bolo baseado no testemunho do vizinho, ou baseada no conhecimento prévio que tenho sobre as habilidades culinárias do vizinho, mas estas não seriam o que estamos chamando aqui de crenças perceptuais. Para estas crenças, preciso estar nesta *relação-de-provar-o-gosto*, preciso executar um conjunto determinado de ações que me coloquem nesta relação particular. Agora, existe alguma relação análoga a *relação-de-provar-o-gosto* quando o objeto de se quer estar consciente é o próprio sujeito que percebe? Existe algo como a *relação-de-sentir-a -si-mesmo*? Ou a *relação-estar-consciente-de-si-mesmo*? Supondo que exista tal relação, faz sentido dizer que eu

possa não estar nesta relação, do mesmo modo que posso não estar na *relação-de-provar-o-gosto* com respeito ao bolo de chocolate do vizinho? A resposta de Shoemaker é negativa:

The obvious fact that there is no organ of introspection is in part the fact that there is no such thing as getting oneself in a position - a position one might not have been in - for making oneself the object of one's awareness. So the failure of introspective self-awareness to satisfy condition (1) is in part its failure to satisfy condition (3).” (Shoemaker 1996, p.210, meu sublinhado)

O que está em jogo no pano de fundo são os temas da discussão de Shoemaker sobre autoreferência (*self-reference*) e autoconsciência (*self-awareness*) que foram tratados em uma série de artigos desde *Self-Knowledge and Self-Awareness* de 1968, passando por *Self-reference and Self-awareness*, do mesmo ano e *Introspection and the Self* de 1986 até ressonâncias destes temas em seu mais recente *Self-Intimation and Second-order Belief*, publicado em 2012 na coletânea *Introspection and Consciousness* (Smithies 2012). Fazer justiça ao detalhamento da discussão de Shoemaker sobre autoconsciência nos levaria para longe dos objetivos desta dissertação, mas não é possível entender seus argumentos contra adequação das condições (3) e (4) para explicar processos introspectivos sem que tenhamos compreendido dois conceitos básicos desenvolvidos por Shoemaker: autoreferência sem identificação (*self-reference without identification*) e imunidade ao erro por identificação equívoca em relação ao pronome da primeira pessoa (*immunity to error through misidentification relative to the first-person pronouns*).

O argumento de Shoemaker contra a adequação da condição de identificação (4) é estritamente dependente destes dois conceitos:

For example, I become aware that George has shaved his beard by seeing George, seeing that he is beardless, and identifying the man I see as George, the man I know to have previously been bearded. So the provision of identification information is an important part of the role played by awareness of objects in giving us awareness of facts. But in introspective self-knowledge there is no room for an identification of oneself, and no need for information on which to base such an identification (see

Shoemaker, 1968 and 1986). There are indeed cases of genuine perceptual knowledge in which awareness of oneself provides identification information, as when noting the features of the man I see in the mirror or on the television monitor tells me that he is myself. But there is no such role for awareness of oneself as an object to play in explaining my introspective knowledge that I am hungry, angry, or alarmed. This comes out in the fact that there is no possibility here of a misidentification; if I have my usual access to my hunger, there is no room for the thought "Someone is hungry all right, but is it me?" (Shoemaker 1996, p. 210-211, meu sublinhado).

Nesta passagem Shoemaker faz menção explícita a seus dois importantes artigos que citamos acima: *Self-Knowledge and Self-awareness*, de 1968 e *Introspection and the Self* de 1986. Na primeira passagem sublinhada, o que está em jogo é o conceito de autoreferência sem identificação. Na segunda passagem sublinhada o conceito envolvido é de imunidade ao erro pela identificação equívoca em relação ao pronome de primeira pessoa. Este conceito é originado da explicação de Shoemaker sobre a distinção que Wittgenstein usa no *Blue Book* entre os usos do pronome da primeira pessoa do singular como objeto e como sujeito. Casos do primeiro uso seriam enunciados como:

- (1) Eu engordei 10 quilos nos últimos 10 meses;
- (2) Eu estou sangrando no braço;
- (3) Eu quebrei minha perna;

Para ficar com os exemplos de Wittgenstein para o segundo uso, podemos citar enunciados como:

- (4) Eu vejo tal-e-tal;
- (5) Eu ouço tal-e-tal;
- (6) Eu tento levantar meu braço;
- (7) Eu penso que vai chover;
- (8) Eu tenho dor de dente

Como Shoemaker alerta, os enunciados (4) a (8) que constituem exemplos de uso do pronome "eu" como sujeito são casos onde os filósofos

estariam tentados a atribuir algum tipo de incorrigibilidade. Mas o ponto de Wittgenstein não é que estes enunciados sejam incorrigíveis em absoluto e sim que eles apresentam imunidade a um tipo de erro em particular, a saber, que não faz sentido perguntar as seguintes contrapartidas:

(4-a) É você mesmo que vê tal-e-tal?

(5-a) Tem certeza que é você mesmo que ouve tal-e-tal?

(6-a) Você tenta levantar seu braço, mas está certo que não é outra pessoa que tenta?

(7-a) É você que pensa que vai chover?

(8-a) Está certo que é você que tem dor de dente?

O fato gramatical de que o conjunto (4-a) a (8-a) seja constituído por perguntas para as quais não possamos dar um uso com sentido, mostra o tipo de imunidade que está envolvida no uso do pronome da primeira pessoa do singular. Ou seja, esta imunidade é derivada de um fato gramatical que exclui perguntas como (4a) - (8a) do jogo de linguagem de autoatribuição de estados mentais, experiências sensório-motoras ou atitudes proposicionais.

Ao contrário, em relação aos enunciados de (1) a (3) podemos imaginar situações onde faça sentido perguntar se é mesmo o meu braço que está sangrando: o exemplo de Wittgenstein pede para imaginar uma cena de acidente onde confundo meu braço ou minha perna com o braço ou a perna de um vizinho ou quando olho a imagem de meu irmão gêmeo no espelho e confundo o reflexo de sua imagem sangrando com a minha situação.

O conceito de autoreferência sem identificação está relacionado com esta mesma discussão, mas por uma via diferente. Ao explicar este conceito em *Self-reference and Self-awareness*, Shoemaker afirma:

My use of the word 'I' as the subject of my statement is not due to my having identified as myself something of which I know, or believe, or wish to say, that the predicate of my statement applies to it. (Brook 2001,p. 84)

O ponto que nos interessa aqui, para compreensão dos argumentos de Shoemaker no contexto da discussão dos modelos perceptuais da introspecção, é o seguinte: o pronome “eu” não identifica sua referência através da descrição de propriedades deste objeto (como seria o caso de uma descrição definida). Da mesma forma, a referência do pronome “eu” não é dada pela intenção do falante, tal como a referência de pronomes demonstrativos como “este” ou “isto”. Sendo assim, a autoconsciência, tomada aqui como o mecanismo pelo qual definimos a referência do pronome pessoal “eu”, não funciona de acordo com a descrição oferecida pela condição (4) – o que é suficiente para o ponto pretendido.

3.1.4 Condição não-relacional

Para análise da adequação da condição não-relacional (5) aos processos introspectivos, Shoemaker retoma a análise do conhecimento dos diversos estados mentais, dividindo seu argumento para consideração em separado dos estados intencionais (crenças, desejos, intenções, expectativas, etc.) e das sensações e experiências sensoriais (dores, arrepios, odores, etc.)

Vejam os primeiro seu argumento para inadequação da condição (5) para percepção introspectiva de estados intencionais:

Condition (5) says that the information about objects we get in perception crucially involves their intrinsic, nonrelational properties. But intentional states are standardly individuated by their contents, and when one knows about one's intentional states introspectively what one knows is, standardly, just that one has a state of a certain kind with a certain intentional content, i.e., that one has a belief that so and so, a desire for such and such, or the like. And recent discussion of mental content seems to have established that a person's having a state with a certain content consists in part in "external" facts about the person's environment - in the person's standing, or having stood, in perceptual relations to external objects of certain kinds, in his belonging to a linguistic community in which certain practices exist, and so on. I have in mind the arguments of Hilary Putnam and Tyler Burge. So having a state with a certain intentional content is not an intrinsic feature of a person, and having a certain content will not be an intrinsic feature of a belief or desire. (Shoemaker 1996,p.212)

Gostaria de reconstruir este argumento de acordo com o seguinte esquema:

- (a) estados intencionais são individuados por seus conteúdos;
- (b) os conteúdos dos estados intencionais de um sujeito S são constituídos, em parte, por relações que S estabelece com fatores externos a S;
- (c) as relações com fatores externos, por definição, não são propriedades intrínsecas de S;
- (d) se fatores externos a S são necessários para definir o conteúdo dos estados intencionais de S [premissa (b)], e se fatores externos, por

definição, não são propriedades intrínsecas de S [premissa (c)], então a individuação dos estados intencionais de S não depende de propriedades intrínsecas de S;

- (e) o conseqüente de (d) nega a condição (5), logo esta condição não se aplica aos estados intencionais de S.

O argumento de Shoemaker depende da aceitação da premissa (b), que equivale a aceitar a tese do externismo sobre conteúdo mental, tal como foi exposta e defendida nos artigos *The meaning of "meaning"* (1975) de Hilary Putnam e *Individualism and the Mental* (1979) de Tyler Burge. Muita tinta tem sido gasta na literatura recente para explorar as implicações do externismo de conteúdo para as teses da autoridade da primeira pessoa. Tyler Burge, John Heil e o próprio Shoemaker estão entre os defensores de uma teoria compatibilista, que aceita o externismo de conteúdo e não nega autoridade de primeira pessoa. Registro estas referências apenas para situar o argumento de Shoemaker no *habitat* que lhe é próprio. Avaliar os erros e acertos do externismo e a relação destes erros e acertos para as teorias contemporâneas sobre autoconhecimento não é nosso objetivo aqui. Acredito, inclusive, que esta discussão adquiriu um volume desproporcional à sua relevância, e que esta discussão já poderia ser dada por encerrada nos termos em que o próprio Shoemaker encaminha a solução em uma nota de rodapé do texto que estamos aqui examinando que, dada sua clareza, cito na íntegra:

It may seem that the claim that the content of mental states is fixed in part by states of affairs outside the head of the subject of the mental states poses a problem for any theory of introspective self-knowledge, and not only for those that invoke the inner sense model. If what makes it the case that my thoughts are about water, and not about the different stuff twater that abounds on Putnam's Twin Earth, is the fact that if it is water rather than twater that abounds in my environment, then how can I know that I am thinking about water without investigating my environment? The short answer to this is that the contents of mental states are fixed holistically, and that whatever fixes the content of the first-order belief I express by saying "There is water in the glass" also fixes in the same way the embedded content in the second-order belief I express by saying "I believe there is water in the glass"-assuming that "water" would be used by me

univocally in those reports. Given that the first is about water rather than twater, the second will ascribe a belief about water rather than a belief about twater. So to explain how I can know what I believe without investigating my environment, all we need to suppose is that I am such that having a first-order belief with a certain content typically gives rise to a second-order belief that one has a belief with that content. In such an account the inner sense model has not so far been invoked, and so far there is no mystery. The mystery arises when we say that the first-order belief gives rise to the second-order belief by means of a process involving a perception of intrinsic, nonrelational features of the first-order belief – for we have no idea what such features could be. (Shoemaker 1996,p. 213, meus sublinhados)

Argumentar na direção da minha concordância com a posição que Shoemaker adota nesta citação é tarefa que delego para outra ocasião. Importante, no momento, é destacar dois pontos: (1) a dependência anafórica para fixação do conteúdo entre (p) e [creio/desejo/espero que (p)]: “o que quer que seja que fixa o conteúdo da crença de primeira ordem que expresso dizendo “Há água no copo” fixa também da mesma maneira o conteúdo contido na crença de segunda ordem que expresso dizendo “Acredito que há água no copo” e (2) para compreender a relação constitutiva, a priori, entre a crença de primeira ordem “Há água no copo” e a crença de segunda ordem “Creio que há água no copo”, “tudo que precisamos supor é que sou de tal maneira que ter a crença de primeira ordem com um certo conteúdo tipicamente dá origem à crença de segunda ordem que têm-se a crença com tal conteúdo.”

Voltando à crítica de Shoemaker, o argumento que citamos e apresentamos na reconstrução acima não é o único ponto de apoio contra a adequação da condição (5) aos processos introspectivos.

Um passo adiante é dado quando pensamos sobre o modo como nos tornamos conscientes sobre um estado intencional ocorrente. Digamos que em uma conversa com amigos eu afirme “Creio que o Inter será campeão brasileiro em 2013”. Após ouvir vários argumentos contrários, mostrando que a probabilidade do Inter tornar-se campeão em 2013 é menor que 1%, reviso minha crença a ponto de abandoná-la. Neste caso passo a reconhecer que

expressaria melhor meu ponto de vista futebolístico dizendo “Desejo que o Inter seja campeão brasileiro em 2013”. É possível que eu estivesse, desde sempre, enganado em relação à atitude proposicional que eu tinha em relação ao conteúdo proposicional [Inter será campeão brasileiro em 2013].

Agora, por mais que eu tivesse confundido a manifestação de uma crença com a expressão de um desejo, nenhum argumento que meus amigos pudessem elencar me convenceria que identifiquei de maneira equívoca a crença ou desejo sobre o desempenho esportivo do Inter com a crença expressa pelo enunciado “Creio que o Grêmio será campeão brasileiro em 2013”. Esse é um dos pontos de Shoemaker na sequência de seu argumento:

But there are other reasons for denying that beliefs and such can be the nonfactual objects of a kind of perception. I am aware that I believe that Boris Yeltsin is President of Russia. It seems clear that it would be utterly wrong to characterize this awareness by saying that at some point I became aware of an entity and identified it, that entity, as a belief that Boris Yeltsin holds that office. To say that would suggest that it ought to be possible for someone to become aware of a belief and misidentify it as something other than a belief, or as a belief with a content other than the one it has; or that it ought to be possible for someone to become aware of something that is not a belief, say a wish that Adlai Stevenson had been elected President of the United States, and misidentify it as a belief that Boris Yeltsin is President of Russia on the basis of the intrinsic features one observes it to have. And while mistakes about one's propositional attitudes are no doubt possible, these kinds of mistakes seem clearly not to be. (Shoemaker 1996, p. 213)

Quando me torno consciente de um estado intencional ocorrente, não é como se eu tivesse encontrado este estado através de uma espécie de escrutínio do que se passa na minha cabeça, identificando-o por propriedades intrínsecas deste estado, da mesma maneira que identifico o trecho de uma música conhecida através da intensidade do timbre ou da harmonia de uma sequência de acordes. Quais seriam as propriedades intrínsecas de uma crença ou de uma expectativa? Não parece sensato pensar que existam estas propriedades intrínsecas, através das quais identificamos crenças, desejos, expectativas, suposições ou quaisquer estados intencionais.

3.1.5 Condição de atenção

Depois de reconhecer que o modelo perceptual ganha plausibilidade e apelo intuitivo quando usado para explicar a consciência de sensações e experiências sensoriais, em particular no caso da dor e das pós-imagens, Shoemaker argumenta em relação a última condição que é especialmente importante para o modelo objeto-percepção; a saber a condição de atenção (6). O que significa dizer que presto atenção em minha crença expressa no enunciado “Creio que o Inter será campeão brasileiro em 2013”? O que significa dizer que alterno minha atenção entre as crenças expressas nos seguintes enunciados: “Creio que o Inter será campeão brasileiro em 2013” e “Creio que o próximo verão será escaldante”? Em nenhuma destas operações da atenção eu tenho uma experiência sensorial particular. No primeiro caso eu penso sobre as possibilidades do Inter ser campeão, como na conversa com meus amigos no exemplo acima. No segundo caso, eu deixo de pensar no Inter e começo a pensar em quão desagradável poderá ser o próximo verão em razão das altas temperaturas que estou esperando que aconteçam. Mas mudanças na direção da atenção quando o que está em jogo são estados sensórios jogam um papel completamente diferente. Posso mudar minha atenção da sensação tátil de meu pé direito dentro do sapato para a sensação de leve desconforto no meu ombro esquerdo. Há nesta operação uma diferença sensorial. No caso de estados intencionais não-sensórios, como crenças, desejos e expectativas, não há nada de perceptual em “deixar de pensar em...” ou “começar a pensar em...”:

...there remains one consideration that makes it more natural to think of our awareness of sensory states in perceptual terms than it is to think of our awareness of beliefs, desires, etc., in such terms. That is the role of *attention* in the two domains. It would seem that one can shift one's attention from one sensation to another, or from one sensory experience to another, in the same way one can shift one's attention from one object of perception to another. But the notion of attention is not really at home in the domain of the nonsensory intentional states such as belief (Shoemaker 1996, p.218, meu sublinhado)

3.1.6 Modelo perceptual amplo e “self-blindness”

O modelo perceptual amplo é caracterizado por duas condições:

(7) Condição Causal:

“Crenças perceptuais são causalmente produzidas pelos objetos ou estados de coisas percebidos, através de um mecanismo causal que normalmente produz crenças que são verdadeiras.”

(8) Condição de independência:

“Os objetos e estados de coisas que são percebidos e sobre os quais a percepção fornece conhecimento, existem independentemente de serem percebidos, e, com certas exceções, independentemente de existirem coisas com a capacidade para percebê-los ou serem deles conscientes.”

Nos primeiros parágrafos da segunda *Royce Lecture*, Shoemaker contrasta o modelo perceptual amplo com o que ele chama de “concepção cartesiana” sobre o conhecimento de nossos estados mentais. De acordo com a “concepção cartesiana”:

...the mind is transparent to itself. It is of the essence of mental entities and mental facts, of whatever kind, to be conscious, where a mental entity’s being conscious involves its revealing its existence and nature to its possessor in an immediate way. This conception involves a strong form of the doctrine that mental entities are “self-intimating”, and usually goes with a strong form of the view that judgments about one’s own mental states are incorrigible and infallible, expressing a super-certain kind of knowledge which is suited for being the epistemological foundation for the rest of what we know. (Shoemaker 1996, p. 224)

O objetivo de Shoemaker nesse momento de sua argumentação é criticar o modelo perceptual amplo sem recair na concepção cartesiana. Se reescrevermos a condição (8) para o caso das crenças introspectivas teremos algo nos seguintes termos:

(8)*: Os objetos e estados de coisas que são introspectados e sobre os quais a introspecção fornece conhecimento, existem independentemente

de serem introspectados, e, com certas exceções, independentemente de existirem coisas com a capacidade para introspectá-los ou serem deles conscientes.

Shoemaker considera uma consequência de (8)* que é ao menos logicamente possível pensar em criaturas que tivessem os conceitos apropriados para descrever determinados estados mentais, mas que fossem sistematicamente incapazes de introspectar estes estados. Nos casos de “blindsight” temos um sujeito que, em função de lesões cerebrais, alega não ter estados visuais conscientes sobre um determinado espaço de seu campo visual (apesar de competência estatística acima da média para dar palpites sobre o que estaria “vendo”). Shoemaker constrói um caso análogo para a introspecção, através da suposição de “self-blindness” ou cegueira introspectiva:

To be self-blind with respect to a certain kind of mental fact or phenomenon, a creature must have the ability to conceive of those facts and phenomena (just as the person who is literally blind will be able to conceive of those states of affairs she is unable to learn about visually). So lower animals who are precluded by their conceptual poverty from having first-person access do not count as self-blind. And it is only introspective access to those phenomena that the creature is supposed to lack; it is not precluded that she could learn of them in the way others might learn of them, i.e., by observing her own behavior, or by discovering facts about her own neurophysiological states. (Shoemaker 1996, p. 226)

Shoemaker argumenta contra a possibilidade de cegueira introspectiva. Caso estabelecida a cogência de seu argumento, teríamos que rejeitar a condição (8)*, e com ela a adequação do modelo perceptual amplo ao caso da introspecção. A partir deste resultado, a posição alternativa adotada por Shoemaker para explicar o tipo de relação entre estados mentais e nosso conhecimento destes estados, que veio ser nomeada como “Constitutivismo”, é apresentada nos seguintes termos:

The view I support holds that there is a conceptual, constitutive connection between the existence of certain sorts of mental entities and their introspective accessibility, while denying the transparency of the mental. It is a version of the view that certain facts are “self-intimating” or “self-presenting”, but a weaker version than the transparency thesis. (Shoemaker 1996, p. 225, meu sublinhado)

Vemos que esta formulação afeta também a contrapartida da condição (7) para crenças introspectivas, já que a relação entre um estado mental e a crença introspectiva não é mais considerada como uma relação causal, logo contingente, tal como defendida por Armstrong em seu argumento das “existências independentes” com vimos em 2.1.4.

Este é o percurso da argumentação de Shoemaker contra o modelo perceptual amplo, mas para aceitar o argumento precisamos primeiro garantir a premissa inicial do argumento, a saber, que “self-blindness” é uma impossibilidade. Shoemaker analisa separadamente a impossibilidade de ser introspectivamente cego em relação a (i) dores, (ii) crenças, (iii) experiências perceptuais e (iv) intenções em ações voluntárias. Vou desconsiderar seus argumentos para (ii), (iii) e (iv) e apresentarei o caso da impossibilidade de cegueira introspectiva para dores. A razão para suprimir a análise dos demais casos é o simples fato de que os argumentos se sobrepõem, utilizando uma estratégia similar que é contrastar a cegueira introspectiva com as pressuposições da agência racional – que é o que nos interessa destacar.

3.1.6 (i) Cegueira introspectiva e dores

Apesar da recomendação de cuidado ao examinar o caso da dor como um estado mental que, dado sua tipicidade, pode induzir a erro, podemos resumir o ponto principal do argumento de Shoemaker no seguinte esquema.

- (1) Um sujeito que têm dores, comporta-se de maneiras determinadas. Procura um médico ou toma algum medicamento, por exemplo.
- (2) Comportamentos típicos de dor pressupõem o desejo que a dor cesse.
- (3) Ter o desejo que a dor cesse, pressupõe acreditar que se está com dor.
- (4) Para um agente cujo comportamento é motivado por razões, acreditar estar com dor é estar consciente de que se está com dor.
- (5) Logo, para ter comportamentos de dor, quando este comportamento é motivado por razões, tenho que estar consciente da dor.

Creio que este esquema captura o essencial do que Shoemaker diz no seguinte trecho:

One thing that seems clear is that at least some kinds of "pain behavior" – taking aspirin and calling the doctor are good examples – are intelligible *as* pain behavior only on the assumption that the subject is aware of pain; for to see them as pain behavior *is* to see them as motivated by such states of the creature as the belief that it is in pain, the desire to be rid of the pain, and the belief that such and such a course of behavior will achieve that result. Indeed, to say that a creature wants to be rid of pain presupposes that it believes that it is in pain. One can want not to have something while being agnostic about whether one has it; but one can't want to be rid of something, to cease having it, So it seems that we cannot suppose that our self-blind creature has a desire to be rid of its pain, or engages in the kinds of pain-behavior that would be motivated by such a desire without believing that one currently does have it. (Shoemaker 1996, p. 228)

O ponto que gostaria de destacar está implícito na premissa (4). Neste contexto, agir por razões não exige nada de muito sofisticado. Posso considerar razões para agir, por exemplo, ter o desejo que a dor cesse e procurar o médico para que me administre algum medicamento para este fim. O essencial do que

formulei como a premissa (4) em nosso esquema é perfeitamente capturado por Brie Gertler:

Shoemaker's conception of rationality is crucial to this argument. For him, the requirements of rationality form a bridge between pain and awareness of pain. Rational creatures are, simply put, creatures capable of reasoning. And reasoning – specifically, reasoning about what to do, so-called “practical” reasoning – involves beliefs and desires. I cannot reason about what I should do next unless I have desires (e.g. to eat) and beliefs (e.g. about which things in my environment are edible). So rational creatures are at least capable of believing and desiring. In a creature capable of believing and desiring, no state will qualify as a *pain* unless that state (usually) causes a *desire* to be rid of it. And one cannot desire to be rid of a pain unless one *believes* that the pain is present. So a rational creature in pain will (usually) believe that she is in pain. (Gertler 2011, p. 150)

O que é crucial para o argumento não é apenas a concepção de Shoemaker sobre racionalidade. O que ele está exigindo – adequação entre comportamento, desejos e crenças – é tão pouco que podemos considerar que esta é uma concepção mínima de racionalidade. A partir da conclusão do argumento podemos inferir que a consciência introspectiva é condição necessária para a agência racional, *simpliciter*. Se além de necessária é condição suficiente não é uma questão para a qual Shoemaker tenha uma resposta nas *Royce Lectures*, mas pelo menos este resultado nos está assegurado:

So it is a necessary condition of one's having knowledge of one's own agency, even inferential knowledge of it, that one not be self-blind with respect to one's belief and desires. (Shoemaker 1996, p. 235)

3.1.7 Constitutivismo, conjunções mooreanas e argumento moral

A tese do constitutivismo defendida por Shoemaker pode ser expressa pela seguinte fórmula:

Def.= seja um estado mental *EM* e um sujeito *S*, *S* está em *EM* se e somente se *S* acredita que está em *EM*.

Formulada para estados mentais de crença, temos que *S* acredita que *p*, se e somente se, *S* acredita que acredita que *p*. Vejamos como a aplicação desta regra é frutífera para explicar a anomalia das conjunções mooreanas: um dos temas caros a Shoemaker.

A discussão sobre o paradoxo de Moore me deixa com a estranha impressão de que estamos às voltas com um falso problema que possui várias soluções. A estranheza nasce disso: se o paradoxo de Moore é um falso problema, nenhuma solução deveria ser possível a não ser a dissolução do problema, mostrando que não se trata, propriamente, de um paradoxo. Independentemente de dar uma resposta à questão se o paradoxo de Moore apresenta um problema genuíno, vou evitar falar de “paradoxo de Moore”. A expressão “conjunções mooreanas”, por esta razão, é mais apropriada e menos comprometedora.

Mas, afinal, qual o problema expresso pela conjunção mooreana padrão? O diagnóstico usual, bem como uma alternativa de solução (aliás, a do próprio Moore) é claramente explicado neste trecho de Sydney Shoemaker:

Moore's paradox is usually presented by pointing to the puzzling character of certain sentences, or of imagined utterances or assertions involving these sentences. Conjunctive sentences such as Max Black's "Oysters are edible, but I don't believe that they are," or the more usual "It is raining but I don't believe that it is raining," I will call "Moore-paradoxical sentences." These are seen as having a logically anomalous character because their assertive utterance would involve some kind of logical impropriety, despite the fact that both of their conjuncts could be true. And I think it is widely assumed that both the paradox and its resolution have to do with the linguistic expression of belief. A natural first move to make in explaining it is to say that while

what someone actually says in uttering a Moore-paradoxical sentence is not self-contradictory, since both conjuncts could be true, there is nevertheless a contradiction between what the speaker in some way implies in asserting the first conjunct, namely that she believes the thing asserted, and what she explicitly says in asserting the second, namely that she does not believe that thing. This is also put by saying that there is a contradiction between something expressed by the first conjunct and what is asserted by the second.” (“*Moore’s paradoxes and self-knowledge*” em Shoemaker 1996, p. 74)

O que seria esta “impropriedade lógica” que não se resume a uma autocontradição? A conjunção mooreana padrão será aqui apresentada como um caso de ponto cego doxástico, a saber, a conjunção mooreana padrão não seria acessível à crença. Este seria seu “caráter logicamente anômalo”. Adotaremos a seguinte definição de ponto cego:

Def. : Uma proposição p é um ponto cego para uma atitude proposicional A e um sujeito S (num dado tempo t) se, e somente se, p é consistente mas S não pode ter a atitude A para com p .

Antecipo que considero a solução apontada pelo próprio Moore como plenamente satisfatória para explicar o que, *prima facie*, nos causa espanto quando nos deparamos com sentenças do tipo “Está chovendo lá fora, mas eu não acredito nisso”. É claramente implicado na afirmação da primeira sentença (“Está chovendo lá fora”) que eu creio no que afirmo. Esta implicação é contraditória com a segunda sentença (“mas eu não acredito nisso”). Logo, contradição. QED.

A regra segundo a qual as pessoas acreditam no que afirmam seria uma regra da asserção genuína. Uma regra conversacional perfeitamente justificada dentro de uma teoria dos atos de fala, aos moldes griceanos. Novamente com a palavra, Shoemaker:

I think that there is a good deal of truth in accounts along these lines, and I do not intend to quibble with their specific claims. However, I do want to question the claim that Moore's paradox has to do solely or primarily with the linguistic expression of beliefs, and that its resolution lies in speech act theory - in considerations about the

nature of assertion, the nature of linguistic intentions, and the like. (Shoemaker 1996, p. 75).

O ponto para o qual Shoemaker está chamando a atenção neste artigo é exatamente o mesmo tema de nossa sugestão no que diz respeito ao tratamento da conjunção mooreana padrão como um ponto cego doxástico.

What seems to me too little noticed is that there is something paradoxical or logically peculiar about idea of someone's believing the propositional content of a Moore-paradoxical sentence, whether or not the person gives linguistic expression to this belief. What really needs to be explained is why someone cannot coherently believe that it is raining and that she doesn't believe that it is, despite the fact that the conjuncts of this belief can both be true. If we can show that such beliefs are impossible, or at least logically defective, and if we can come up with an explanation of this, then an explanation of why one cannot (coherently) assert a Moore-paradoxical sentence will come along for free, via the principle that what can be (coherently) believed constrains what one can be (coherently) asserted. (Shoemaker 1996, p. 75-6, grifo meu)

A explicação mais sucinta que conheço para este ponto é oferecida por Jaakko Hintikka em seu sistema de Lógica Epistêmica (Hintikka 1962). Hintikka oferece uma solução para o que ele chama de “problema de Moore de dizer e não crer” mostrando que a seguinte proposição (A) gera uma contradição, ou, nos termos de Hintikka, a proposição (A) é doxasticamente indefensável.

(A) “ $Ba (p \wedge \neg Bap)$ ”³

A prova sugerida é bastante simples:

Supondo que:

“ $Ba (p \wedge \neg Bap)$ ” $\in \mu$

- | | |
|--|---------------------------------------|
| (1) “ $p \wedge \neg Bap$ ” $\in \mu^+$ | da hipótese, por (C.b ⁺); |
| (2) “ $Ba (p \wedge \neg Bap)$ ” $\in \mu^+$ | da hipótese, por (C.BB ⁺) |

³ A expressão Bap lê-se “O sujeito a crê que p ”.

- | | |
|--|-----------------------------|
| (3) " $\neg Bap$ " $\in \mu^+$ | de (1), por (C. \wedge); |
| (4) " $Ca \neg p$ " $\in \mu^+$ | de (3), por (C. $\neg B$); |
| (5) " $\neg p$ " $\in \mu^{++}$ | de (4), por (C.C $^+$); |
| (6) " $p \wedge \neg Bap$ " $\in \mu^{++}$ | de (2), por (C.B $^+$); |
| (7) " p " $\in \mu^{++}$ | de (6), por (C. \wedge). |

(5) e (7) contradizem uma regra básica do sistema de lógica epistêmica :

(C. \neg): Se $p \in \mu$, então não é o caso que " $\neg p$ " $\in \mu$.

As demais condições que são pressupostas na estratégia redutiva usada por Hintikka tampouco parecem expressar alguma tese contestável. Vejamos:

(C.b $^+$): Se " Bap " $\in \mu$ e se μ pertence a um sistema modelo Ω , então há em Ω ao menos uma alternativa epistêmica μ^+ a μ (com respeito a a) tal que $p \in \mu^+$;

(C.BB $^+$): Se $Bap \in \mu$ e se μ^+ é uma alternativa epistêmica a μ (com respeito a a) em algum sistema modelo, então " Bap " $\in \mu^+$

(C. \wedge): Se " $p \wedge q$ " $\in \mu$ então $p \in \mu$ e $q \in \mu$

(C. $\neg B$): Se " $\neg Bap$ " $\in \mu$, então " $Pa \neg p$ " $\in \mu$

(C.C $^+$): Se " Cap " $\in \mu$ e se μ^+ é uma alternativa epistêmica a μ (com respeito a a) em algum sistema modelo, então $p \in \mu^+$ (" Cap " lê-se como: É compatível, para tudo que a sabe, que p)

(C.B $^+$): Se " Bap " $\in \mu$ e se μ^+ é uma alternativa epistêmica a μ (com respeito a a) em algum sistema modelo, então $p \in \mu^+$

Creio, pelo exposto acima, que o caráter anômalo da conjunção mooreana possa ser descrito em termos de indefensabilidade doxástica. No entanto, é sabido que parte da perplexidade despertada pela conjunção mooreana padrão é que a proposição (A1) não é autocontraditória.

(A1) ($p \wedge \neg Bap$)

Aliás, é uma condição da definição de ponto cego doxástico que a proposição a ser acreditada seja consistente. Para considerar a conjunção mooreana padrão como um caso de ponto cego, (A1) deve ser consistente.

A primeira diferença que salta aos olhos entre (A1) e as típicas autocontradições explícitas (como “ $p \wedge \neg p$ ”) é que modificações de tempo verbal e pessoa fazem desaparecer o caráter anômalo de (A1).

(A2): “ p ” mas ele não acredita que “ p ”

(A3): “ p ” mas eu não acreditava que “ p ”

Estas sentenças não apresentam nenhum problema.

Temos um caso de uso peculiar do verbo “crer” que também normaliza conjunções entre dizer e descrever:

(A4) Maria me traiu, mas eu não acredito que ela tenha feito isso.

Neste caso, a expressão “mas eu não acredito que ela tenha feito isso” é usada no sentido de uma reprimenda, significando algo do tipo “me surpreende que ela tenha feito isso”. Não lemos “eu não acredito ...” como contraditório com a afirmação da traição de Maria. Imagine uma outra cena: você sai de casa para uma rápida ida ao açougue e quando volta a sua casa, seu labrador destroçou o sofá da sala. Você exclama, irado: “Não acredito que ele fez isso!”. Estes casos, em absoluto, apresentam usos do verbo crer que possam formar uma conjunção mooreana, mesmo que usando as mesmas expressões da linguagem ordinária.

(A4) guarda uma similaridade com os exemplos que Richard Moran (2001, p. 71) oferece de verbos psicológicos que, em determinados usos, não fazem referência ao estado mental do falante, mas cumprem outras funções:

(B) Temo ser necessário pedir para que você saia;

(C) Pensei ter dito para que esperasse no carro;

Assim como em (A4) “eu não acredito que...” não significa que o falante não creia que Maria o traiu, em (B) “Temo” não descreve um medo e em (C) “Pensei” não descreve um pensamento que o falante teve. Proposições como (A4), (B) e (C) onde são empregados verbos tidos como “psicológicos” mas que não fazem referência aos estados mentais dos falantes, são uma das vias de normalização da conjunção mooreana padrão. De acordo com esta interpretação, que Moran chama de “visão apresentacional”, o emprego do verbo “eu acredito que (*p*)” em sentenças na primeira pessoa do tempo presente do modo indicativo indicaria apenas um modo de apresentar o conteúdo proposicional:

That is, when someone says, “I think it’s raining out”, his statement does not refer to his (or anyone else’s) state of mind, but is instead simply a more guarded way of making the assertion about the rain. So in this context the word “believe” is not operating as a psychological verb at all. On this view, then, the two parts of the Moore-type statement do not in fact have different subject matters (one part about the rain, and hence the other about someone’s belief); instead, they’re both about the rain, and hence their conjunction really does form a contradiction. And so, what is puzzling about the original statement reduces to the fact that it is a contradiction after all, but in a disguised form. (Moran 2001, p. 71)

Quando afirmei que o paradoxo de Moore me parecia um falso problema com várias soluções, estava pensando em alternativas como essa da “visão presentacional”. Parece sensato pensar que, em certos contextos, “Eu creio que (*p*)” não é uma asserção sobre meu estado mental e sim uma maneira mais “reservada” de afirmar o conteúdo proposicional de (*p*) Sendo assim, isto explicaria a estranheza das seguintes proposições:

(D) Creio que $121 + 15 = 399$

(E) Creio que $2 + 2 = 4$

Se o verbo crer na primeira pessoa do singular, no tempo presente, modo indicativo, é uma forma mais reservada de apresentar o conteúdo proposicional de (*p*), então não faria sentido afirmar (D) ou (E) – o que parece ser o caso. Não

parece fazer sentido afirmar “de modo reservado” que $2 + 2 = 4$. O problema aqui não está na atitude proposicional e sim no conteúdo proposicional.

A “visão apresentacional”, apesar de apontar para algo importante, nos deixa com algumas promissórias de difícil resgate. Em especial, temos que explicar por qual razão, apenas no tempo presente do modo indicativo e na primeira pessoa do singular, a expressão “creio” em “Creio que vai chover amanhã” modula a asserção ao invés de referir a um estado mental do falante.

(F) Eu acreditava que choveria hoje.

(G) Ele acredita que choverá hoje.

Em (F) e (G) o verbo “crer” recupera sua função de referir a um estado mental. Outra das dificuldades dessa proposta é explicar a peculiaridade do verbo “crer” modulando a asserção, em comparação com o emprego de outros verbos psicológicos, no mesmo tempo verbal, modo e conjugação.

(H) Eu suponho que vá chover hoje.

(I) Eu espero que chova hoje.

(J) Eu desejo que chova hoje.

(K) Eu temo que vá chover hoje.

Nos casos de (H) a (K) é trivialmente verdadeiro que os verbos empregados fazem referência a um estado mental do sujeito.

Vamos pensar agora em um cenário para a conjunção mooreana padrão. O objetivo é ver o problema de um ponto de vista completamente distinto. Suponhamos que você me diga: “Não acredito que vai chover esta semana” e, ato contínuo, saia de casa em direção ao trabalho com um guarda-chuva, vestindo uma capa de chuva. O que devo pensar dessa cena? A minha primeira reação é que você está mentindo para mim. Mas suponhamos que o personagem desta história seja alguém que sempre faz suas asserções de forma sincera. Aliás, esta condição é sempre necessária para que as conjunções

mooreanas tenham efeito, a saber, que a asserção seja sincera. Levando isto em conta, minha segunda reação é pensar que você está agindo de uma forma estranha. Há uma clara descontinuidade entre o que você afirmou e o que você fez. Ter uma crença significa, entre outras coisas, agir de uma determinada maneira que seja compatível com a crença asserida. A conjunção mooreana apresenta uma espécie de descontinuidade na agência. Não preciso supor ficções wittgensteinianas de dois sujeitos falando pela minha boca para saber que algo vai mal com o personagem de nossa história. E o sintoma desta estranheza neste caso é análogo ao caso padrão de afirmar “Está chovendo, mas não creio que está chovendo.” Neste sentido, é pertinente associar a explicação da anormalidade das conjunções mooreanas com fenômenos como a *akrasia* e o autoengano, tal como sugere Moran:

An examination of the first-person should account for why someone’s need to rely on behavioral evidence to report on his mental states would suggest something wrong with him, some state of dissociation, and would raise doubts about the rationality of those attitudes of his which are not accessible to him in the normal “immediate” way. In its most radical form, the sort of alienation we’re considering would place the person in a situation he could only describe in a utterance that was a version of Moore’s Paradox: I know this must be envy that I feel toward him (I believe it’s raining out), although there is nothing about him to be envied (but it’s not raining out). (Moran 2001 ,p 69).

Neste ponto podemos perceber como a tese constitutivista é relevante para uma imagem da agência moral. Assim como é estranho contar com evidência sobre meu comportamento para descrever meus estados mentais, podemos considerar que estados mentais como os descritos nos fenômenos da *akrasia* e do autoengano são disfuncionais, pelo simples fato de apresentarem pontos cegos doxásticos, ou seja, apresentam estados mentais com anomalias semelhantes ao caso representado pelas conjunções mooreanas. Do ponto de vista da primeira pessoa, que um estado mental seja acessível à crença não é apenas uma condição necessária para considerá-lo como um estado mental *tout court*. Mais do que isso, a condição de acesso doxástico é uma regra da agência moral autoconsciente. Formulada para estados mentais que expressam intenções ou desejos a tese constitutivista teria a seguinte expressão:

Def.= S deseja que p, se e somente se, S acredita que deseja que p.

Def.=S intenciona que p, se e somente se, S acredita que intenciona que p.

Críticos da tese constitutivista apressariam-se para apresentar como contra-exemplos estados mentais que não covariam com a crença do próprio sujeito sobre estes estados. Os casos típicos de autoengano e confabulação são os mais evidentes. Situações de pacientes em tratamento psicanalítico envolvendo negação de determinados comportamentos também apresentam desafios para a tese constitutivista. No entanto, não podemos perder de vista que a exigência expressa na tese resguarda uma noção cara e necessária para o agir moral, a saber, a racionalidade do agente. E, como vimos, o que Shoemaker exige do agente para ser dito racional não é nada mais do que o alinhamento entre comportamento, desejos e crenças. Exigir que a acessibilidade doxástica seja uma marca da intencionalidade de um estado mental é uma forma de deslocar a discussão sobre a epistemologia da introspecção para o campo da psicologia moral. E esta passagem, dentro do marco racionalista das teorias da introspecção, teve seu caminho aberto por Shoemaker .⁴

⁴ Boa parte dos temas tratados nesta seção me foi sugerida em discussões com Paulo Faria no Seminário livre de Filosofia II – Proposições epistemicamente inacessíveis – ministrado em 2011/02.

Capítulo 4 - Retratos da interioridade e conceitos morais

A premissa básica da psicologia moral é que ações morais são ações voluntárias. Esta afirmação está suficientemente enraizada na discussão filosófica sobre a moral e creio que é indiscutível. A correlação conceitual entre ação moral e ação voluntária é suficientemente clara e aparece em momentos inaugurais da tradição do pensamento moral, como na *Ética a Nicômano*, por exemplo. Para discutir as características da ação voluntária em contextos morais torna-se necessária, portanto, a introdução de um conjunto bastante específico de conceitos psicológicos. “Intenção”, “motivo”, “desejo” e “impulso” são alguns destes conceitos. Sendo assim, podemos pensar que uma determinada imagem do mental está sempre presente no pano de fundo da investigação moral. Neste sentido falamos de uma psicologia moral dos estoicos em contraposição a uma psicologia moral aristotélica, platônica, lockiana, shaftesburyana, kantiana e assim por diante, querendo com isto marcar que diferentes teorias sobre o funcionamento dos conceitos psicológicos certamente influenciaram a maneira como determinados autores ou tradições pensaram sobre a moral. Uma das questões que deram origem a esta dissertação é a seguinte: de que maneira as diferentes teorias sobre a introspecção influenciaram o discurso sobre a moral?

Para responder a esta pergunta, ou pelo menos indicar alguns elementos para a resposta, precisamos nivelar o terreno eliminando alguns obstáculos. O primeiro obstáculo seria desafiar a própria pressuposição de que exista alguma relação entre conceitos morais e conceitos psicológicos. Em outras palavras, seria questionar a própria viabilidade do que entendemos como psicologia moral. Este obstáculo me parece artificial. É no mínimo exótico que seja possível negar a relação de intersecção entre o conjunto das ações morais e o conjunto das ações voluntárias. O segundo obstáculo seria dizer que o conceito de introspecção tem uma história muito mais recente do que os conceitos morais. Em sentido estrito este obstáculo é intransponível e teríamos que reformular

nossa pergunta, tornando-a mais elástica. Deixamos de nos perguntar sobre a relação entre os conceitos morais e as diferentes teorias da introspecção e sim sobre a relação destes conceitos morais com o modo como descrevemos o conhecimento que possuímos, alegadamente, sobre nossos estados e processos mentais. Temos uma razão histórica para considerar este obstáculo, *prima facie*, como legítimo:

It has been argued that the ancient Greeks did not have any concept of introspection, or at least the Greeks up to and including Aristotle, though there is a section in *De Anima* where Aristotle may be saying that we can reflect on our own mental states. The earliest account of introspection that could lay claim to be a classical account – though neither the term “introspection” nor any equivalent technical term is employed – is probably to be found in Augustine’s *De Trinitate*. He there remarks, “What then can be the purport of the injunction, Know thyself? I suppose it is the mind should reflect upon itself. (Lyons 1986, p. 1).

Willan Lyons lembra que, de acordo com *A New English Dictionary on Historical Principle*, a primeira ocorrência do termo “introspection” aparece na obra *The Primitive Origination of Mankind, Considered and Examined According to the Light of Nature* de Sir Matthew Hale, editado em Londres no ano de 1677, tornando anacrônico procurarmos por teorias da introspecção anteriores ao século XVII.

Mesmo que tenhamos que refazer nossa pergunta atendo-nos às variedades da autoreflexão sobre estados e processos mentais, seja nas diferentes formulações do mais conhecido entre os preceitos delficos, ou nas diversas maneiras que a tradição filosófica deu expressão ao fato de que autoatribuímos predicados mentais; o ponto de fuga que não podemos perder de vista é o seguinte: diferentes retratos da interioridade sugerem nuances na coloração de nossos conceitos morais ordinários. Para ilustrar o ponto, adotaremos a estratégia de descrever em detalhe alguns destes retratos. A tentativa mais bem sucedida na exposição destas nuances é a obra generosamente erudita de Charles Taylor, *Sources of the Self – The Making of Modern Identity*, que nos servirá de pano de fundo em alguns momentos.

As teorias da introspecção que aproximam-se de modelos perceptuais (ver capítulo 2), privilegiando a visão dentre os recursos sensoriais como parâmetro comparativo, mantiveram-se como a opinião predominante por um longo período. Traços deste tipo de teoria estão na base da analogia lockiana entre sentido externo e interno, que informou grande parte da tradição filosófica posterior. A crítica deste modelo é algo recente, podemos afirmar, sem medo de errar, que os argumentos mais acabados contra os modelos perceptuais da introspecção encontram-se, como vimos no capítulo 3.1, em Sydney Shoemaker. Sendo assim, não temos mais do que 30 anos de discussão sobre este tema. Mesmo se considerarmos a publicação de *The Concept of Mind* em 1949, como um dos textos seminais que lançaram dúvidas sobre a concepção predominante sobre os processos introspectivos, mesmo assim ainda teríamos uma certidão de nascimento relativamente próxima para a discussão filosófica sobre estes temas. Como resultado, é esperado que não tivéssemos formada uma ortodoxia sobre quais teorias da introspecção são mais apropriadas.

Da perspectiva em que vejo o estado atual da discussão, creio que os modelos perceptuais da introspecção estão, em definitivo, fora do jogo. Os argumentos de Shoemaker em "*Self-knowledge and 'Inner Sense'*" parecem-me suficientemente poderosos para afastar qualquer possibilidade de ressuscitar teorias do sentido interno que tomem a analogia da introspecção com os processos sensoriais como ponto de partida. Agora, é claro que nenhuma teoria que tenha sido hegemônica por tanto tempo, deixa o campo de batalha em silêncio. Ressonâncias ainda se fazem sentir de variadas formas e algumas expressões de nossa linguagem ordinária trazem cristalizadas algumas destas imagens que associam introspecção com algum tipo de visão interior. "Olhe para dentro de si", "No fundo ele é uma pessoa diferente", "Procure descobrir sua verdadeira vocação", "Quero ser honesto para comigo mesmo" são apenas algumas destas expressões que usamos de forma corriqueira.

No campo da argumentação moral, reflexos das teorias do sentido interno sobre a introspecção podem ser percebidas claramente. O mais

persistente destes reflexos é o próprio conceito de consciência moral, como uma faculdade interna ao sujeito. Não pretendo fazer aqui uma genealogia conceitual da consciência moral, nem mesmo uma genealogia genética da moralidade aos moldes de Freud ou Nietzsche. Meu objetivo é bem mais modesto. Vou tentar elencar algumas características do que chamarei de visão não-relacional do argumento moral e mostrar como estas características estão vinculadas a propriedades que as teorias do sentido interno atribuem à introspecção. Considerando as deficiências insanáveis das teorias do sentido interno, esta visão não-relacional do argumento moral estaria, por sua vez, viciada na origem.

4.1 Visão não-relacional do argumento moral

Por argumento moral vou entender todas as situações onde deliberamos sobre a propriedade ou impropriedade de uma determinada ação, seja esta realizada por outros ou por nós mesmos. Deliberamos sobre diferentes cursos de ações possíveis, sejam ações que nós realizamos isoladamente ou em conjunto com outras pessoas ou mesmo ações que outras pessoas realizam independentemente de nossa participação. O que estou chamando aqui de argumento moral não diz respeito apenas às ações que pretendemos ou intencionamos realizar; diz respeito também à avaliação que fazemos da propriedade ou impropriedade de ações já realizadas por nós, isoladamente ou em conjunto com outras pessoas, ou já realizadas por outras pessoas sem nossa participação. Grosso modo, o resultado de um argumento moral é uma avaliação que fazemos sobre a correção de uma determinada ação.

Uma visão não-relacional do argumento ou avaliação moral nos diria que os processos pelos quais chegamos a esta avaliação são exclusivos da perspectiva da primeira pessoa, ou seja, não pressupõem que tenhamos que assumir a perspectiva de terceiros ou mesmo levar em conta a opinião destes terceiros em um cenário de conversa, por exemplo.

A imagem que estou tentando retratar é da consciência moral como uma espécie de voz ou luz interior. Ryle tem uma sugestão audaz sobre este ponto:

When the epistemologists' concept of consciousness first became popular, it seems to have been in part a transformed application of the Protestant notion of conscience. The Protestants had to hold that a man could know the moral state of his soul and the wishes of God without the aid of confessors and scholars; they spoke therefore of the God-given 'light' of private conscience. (Ryle 1949, p. 159)

O que Ryle está dizendo aqui é que o conceito de consciência utilizado na epistemologia (e aqui Ryle deve estar pensando em Descartes como exemplo de epistemólogo) é uma transformação do conceito de consciência moral desenvolvido por Lutero e Calvino. Certamente eu não compraria esta primeira

afirmação de Ryle pelo seu valor face. Se ele está referindo-se a Descartes como o epistemólogo em questão, fica ainda mais difícil de aceitar que os conceitos cartesianos de *conscientia* e *sensus internus*, expressos em suas formas latinas, tivessem alguma filiação conceitual com doutrinas teológicas seja de Lutero ou Calvino. No exame mais atento que conheço sobre o uso que Descartes faz destes conceitos não aparece nenhuma filiação protestante para os mesmos. Gordon Baker e Katherine J. Morris no fascinante livro *Descartes' Dualism* (1996) ressaltam antes os aspectos de continuidade entre as definições escolásticas destes termos e as inovações propriamente cartesianas e, em especial, o quão longe a noção de *conscientia* está do conceito moderno de introspecção, tal como o utiliza o próprio Ryle.

Deixando estas discussões de lado, o que me interessa é a segunda parte da afirmação de Ryle quando ele afirma que “um homem poderia conhecer o estado moral de sua alma (...) sem ajuda de confessores ou professores.” (meu sublinhado). Há um ponto nesta sugestão de Ryle que merece ser destacado. Creio que é possível delinear uma tradição cristã de pensamento sobre o tema da consciência moral onde o sujeito moral está como que isolado do convívio de outras pessoas – esta é uma das características do que estou chamando aqui de visão não-relacional, a saber, a ausência de relação com outros agentes morais. Estou pensando na tradição do autoexame de consciência tal como foi desenhado nos *Exercícios Espirituais* de Santo Inácio de Loyola ou nas prédicas de São Boaventura, por exemplo.

Minha sugestão é que a análise da noção cristã de “exame de consciência” ilustra a consideração do que seria uma perspectiva irreduzível em primeira pessoa. Esta noção é expressa de forma lapidar em um texto atribuído a São Boaventura (1221-1274) chamado “*Vida Perfeita para Religiosas*” – e aqui peço desculpas antecipadamente por citar um autor e um texto específico deste autor que destoam das referências bibliográficas típicas das discussões de filosofia moral. Obviamente não vai me interessar aqui saber quão eficiente era

São Boaventura como conselheiro para suas irmãs na fé e sim tentar fixar de modo paradigmático esta noção de “exame de consciência”.

Assim começa São Boaventura o primeiro capítulo desta obra, que leva o subtítulo “*Do verdadeiro conhecimento de si mesmo*”:

Em primeiro lugar, a esposa de Cristo que deseja subir ao ápice da perfeição deve começar por fixar atenção em si, de forma que, esquecida de todo o exterior, entre no segredo de sua consciência, e ali, com diligente cuidado, investigue, examine e veja todos seus defeitos, todos seus hábitos, todas suas fixações, todas suas obras, todos seus pecados, tanto passados como presentes...

As expressões centrais para os meus propósitos são: a) fixar atenção em si, b) esquecida de todo exterior e c) segredo de sua consciência.

Vamos abstrair do fato que São Boaventura está dando conselhos a suas irmãs de fé. Pensemos nestas recomendações como indicações sobre o ponto de vista que devemos adotar para avaliar a conveniência de nossas ações. A imagem que nos fica ao ler esta passagem é que a consciência do agente moral é um lugar para o qual dirijo meu olhar, abstração feita de tudo que não compartilhe do caráter privado desta jornada. Devo me esquecer do que é externo aos segredos da minha privacidade para ter autoconhecimento. Esta imagem não é exclusiva de São Boaventura. Santo Inácio de Loyola (1491 -1556) em suas observações preliminares nos *Exercícios Espirituais* recomenda uma posição de isolamento:

Se alguém tem o tempo mais livre e deseja aproveitar o mais possível, dêem-se a esse exercitante todos os Exercícios espirituais, guardando com exatidão a ordem aqui seguida. Neles, de ordinário, tanto mais se aproveitará, quanto mais o exercitante se separar de todos os amigos e conhecidos, e de todas as ocupações temporais. Mudando-se, por exemplo, da casa onde mora e escolhendo outra, ou outro quarto, para nele viver o mais retirado possível,...

Sempre é bom lembrar a similaridade entre os Exercícios Espirituais de Loyola e a configuração do cenário das meditações cartesianas:

Agora, pois, que meu espírito está livre de todos os cuidados, e que consegui um repouso assegurado numa pacífica solidão, aplicar-me-ei seriamente e com liberdade em destruir em geral todas as minhas antigas opiniões. (Descartes 1641/1973, p.118)

Que o cenário da investigação moral pressuposto na noção de exame de consciência e o cenário da investigação epistemológica cartesiana sejam, ambos, cenários de solidão, de isolamento não parece ser uma coincidência. Neste ponto a sugestão de Ryle de uma proximidade entre a noção de consciência do início da epistemologia moderna e certas concepções religiosas parece não só fazer sentido mas carregar, inclusive, certa plausibilidade histórica.

Se pensarmos na psicologia do agente moral, refletindo sobre seus motivos e intenções, alheio à existência das demandas que outras pessoas colocam sobre seu comportamento, abstraindo até mesmo da existência de outras pessoas, ele permanece isolado no horizonte da reflexão sobre seus interesses e na satisfação do amor-próprio. Quando o espaço da reflexão moral fica restrito à dimensão interna à consciência do sujeito que age, o que sobra, aparentemente, são apenas as considerações do amor próprio. Como afirma o Bispo Joseph Butler (1692-1752), em seu sermão *Upon the Love of our Neighbour*

Let it be allowed, though virtue or moral rectitude does indeed consist in affection to and pursuit of what is right and good, as such, yet when we sit down in a cool hour, we can neither justify to ourselves this [virtue or moral rectitude] or any other pursuit, till we are convinced that it will be for our happiness, or at least not contrary to it. (Butler 1726/2006, p.117)

Recorrer ao Bispo Joseph Butler neste momento da exposição não é acidental. Colocá-lo em uma linhagem comum que passa por Boaventura e Santo Inácio de Loyola é consequente com sugestão feita por Douglas Langston em *Conscience and Other Virtues* (2001) de que o Bispo Joseph Butler talvez seja o principal responsável na tradição filosófica moderna por considerar

consciência - aqui entendido como consciência moral e não como auto-consciência - como uma faculdade, de par com o entendimento e a vontade.

Butler não apenas considera a consciência moral como uma faculdade, mas, além disso, atribui à consciência características de infalibilidade quanto ao julgamento sobre os agentes e suas ações. Cito Douglas Langston:

The authority of conscience over the parts of a human being should be absolute. Any failure to follow the directions of conscience is, in fact, a failure to follow one's own nature. Butler is not shy in attributing qualities to conscience. He regards it as an unerring faculty that judges actions as well as the agent; it is essentially a law unto itself, and it always leads us to the good (37, 43, 45, 69). Its one defect seems to be that it cannot enforce its decisions for, although possessing authority, it lacks power. But there is no doubt in Butler's mind that the world would be better off if conscience had such power: "Had it strength, as it has right; had it power, as it has manifest authority, it would absolutely govern the world". The strong view of conscience one finds in Butler's writing is a departure from the more limited conception found in scholastic writers. It is, however, reminiscent of Luther's views about conscience. Butler, of course, goes beyond even Luther's views, seeing conscience as an unerring faculty. He was apparently influenced by William Perkins on this issue. Perkins considered conscience as the vice-gerent of God, so that the voice of conscience was the infallible voice of God. Similarly, Butler sees conscience as implanted by God to represent God's own commands. (Langston 2001, p.81-2)

O segundo aspecto que caracteriza a visão não-relacional do argumento moral também apresenta uma ressonância protestante. Quando estou isolado no privado da minha consciência ganha importância a reflexão sobre os motivos da ação. Ganha especial destaque a pergunta "Qual a razão para ter feito isso ou aquilo?" A pergunta sobre os motivos da ação em uma visão não-relacional é moralmente mais relevante do que a pergunta sobre as consequências da ação, sejam estas consequências previstas ou não.

A primeira característica da visão não-relacional do argumento moral era a ausência de outros agentes morais no cenário da reflexão, agora temos a minimização das consequências da ação em favor da constante avaliação das

minhas intenções ao agir. Para esta característica podemos pensar em uma filiação calvinista, em razão da obsessão pelo constante autoexame a que a doutrina da eleição incondicional submete os fiéis. As boas intenções seriam sinal da graça, um indício de eleição para salvação divina. Podemos pensar, igualmente, em um cenário menos teologicamente carregado, para buscar filiações históricas para esta característica da visão não-relacional. Creio defensável afirmar que a raiz desta ideia está presente na psicologia moral do estoicismo e, de modo mais genérico, na filosofia moral greco-romana como atesta Pierre Hadot em *Exercices Spirituels et Philosophie Antique*, ao afirmar que a tradição dos exercícios espirituais, tal como a apresentamos em São Boaventura e Ignácio de Loyola, pode ser considerada uma versão católica desta imagem do argumento moral nascida filosofia greco-romana. Quando, na análise de Boaventura, nos deparamos com a expressão “esquecida de todo exterior”, podemos pensá-la de acordo com a orientação de desconsiderar as consequências de nossos atos, fixando-nos, em primeira instância, nas intenções que motivaram minha ação.

Não quero propor uma espécie de esquizofrenia da agência moral, opondo intenções a consequências, de forma absoluta. Parece sensato que uma imagem adequada do argumento moral deva contemplar todos os ângulos da agência. Apenas quero destacar que a visão não-relacional dá prioridade ao exame das intenções. O ponto estoico aqui presente pode ser ilustrado na primeira sentença do *Encheirídion* de Epicteto:

On the one hand, there are things that are in our power, whereas other things are not in our power. In our power are opinion, impulse, desire, aversion, and, in a word, whatever is our own doing. Things not in our power include our body, our possessions, our reputations, our status, and, in a word, whatever is not our own doing. (Seddon 2005, p. 31)

Nossas intenções estão em nosso poder, as consequências de nossas ações, não. O vasto e imprevisível espaço das contingências da causalidade natural está interposto entre o que pretendemos fazer e o que efetivamente

fazemos. As consequências, previstas ou não, na maior parte das vezes, escapam ao controle da agência racional. Por esta razão, a sentença de Epicteto coloca do lado das coisas que estão em nosso poder os estados mentais como opinião (*hupolêpsis*), impulso (*hormê*), desejo (*orexis*) e aversão (*ekklisis*). É com isso que o argumento moral, de acordo com a visão não-relacional, deve estar prioritariamente preocupado. Epicteto vai um pouco mais longe e afirma que proceder de outra forma é dar espaço ao sofrimento. É tarefa para outro momento detalhar a psicologia moral estoica e, em especial, este otimismo sobre estados mentais que estão sob nosso controle que marca indelevelmente a tradição do pensamento moral.

A terceira característica que me parece originada nas teorias perceptuais da introspecção é o caráter pretensamente descritivo dos conceitos morais envolvidos na agência. No início deste capítulo afirmamos que para entender a ação voluntária em contextos morais precisamos introduzir um conjunto específico de conceitos psicológicos, conjunto este que intersecciona a lista de estados mentais de Epicteto.

As teorias perceptuais da introspecção favorecem a concepção segundo a qual conceitos como “intenção”, “motivo”, “desejo”, “impulso” e “aversão”, por exemplo, descrevem estados mentais ocorrentes que podem ser percebidos internamente (Brentano), escaneados por sistemas de autodeteção (Lycan), ou descobertos quando olhamos para dentro de nossas mentes (William James) ou, ainda, quando nos tornamos conscientes de acontecimentos mentais específicos (Armstrong). O sentimentalismo moral, em alguma medida, parece depender da pressuposição de que estes conceitos descrevem estados mentais ocorrentes que são percebidos contingentemente. Já vimos nos comentários sobre Moran o quanto as teorias racionalistas pressionam esta tese, afirmando que as peculiaridades da perspectiva da primeira pessoa são melhor capturadas quando valorizamos o fato de moldarmos nossas intenções e impulsos ao invés de apenas descobri-las através de um pretense autoexame, ou quando

constituímos nossas aversões ou motivos ao invés de observá-los com o “olho da mente”.

A ideia que a introspecção, concebida de acordo com o modelo perceptual, devesse servir como parâmetro para uma correta descrição dos procedimentos de avaliação da ação, ao menos em um sentido bem estrito, onde do ponto de vista do sujeito a questão sobre a correção de um curso de ação se põe tão somente quando faz sentido perguntar quais foram meus motivos ou intenções para fazer o que fiz, está diretamente relacionada com esta terceira característica da visão não-relacional. Se dar razões é o começo de qualquer explicação sobre a responsabilidade pelo que faço ou deixo de fazer, deve ser possível para o agente poder descrever de maneira inequívoca os motivos de suas ações. Esta ideia parece regular uma noção básica de responsabilidade.

A quarta e última característica da visão não-relacional que quero destacar é a privacidade destes estados mentais intencionais. Parece-me correto afirmar que essa concepção da introspecção é o correlato epistemológico de uma certa ideia de que a avaliação da correção moral de minhas ações deveria, em última instância, ser comparada com minhas intenções e motivos – os quais apenas eu poderia descrever corretamente. Este ponto me foi sugerido pela leitura da seguinte passagem de J.L. Austin:

In the case of promising- for example “ I promise to be there tomorrow” – it’s very easy to think that the utterance is simply the outward and visible (that is, verbal) sign of the performance of some inward spiritual act of promising , and this view has certainly been expressed in many classic places. There is the case of Euripide’s Hippolytus, who said “My tongue swore to, but my heart did not” – perhaps it should be “mind” or “spirit” rather than “heart” but at any rate some kind of backstage artiste. Now it is clear from this sort of example that, if we slip into thinking that such utterances are reports, true or false, of the performance of inward and spiritual acts, we open a loophole to perjurers and welschers and bigamists and so on, so that there are disadvantages in being excessively solemn in this way. It is better, perhaps, so stick to the old saying that our word is our bond. (Austin 1979, p.236, meu sublinhado)

Em que sentido, perguntar sobre as “verdadeiras intenções” do sujeito que promete, no exemplo de Austin, é não somente irrelevante mas também, em certo sentido, pernicioso? Em outro momento Austin repete com mais força o mesmo ponto:

It is gratifying to observe in this very example how excess of profundity, or rather solemnity, at once paves the way for immorality. For one who says “promising is not merely a matter of uttering words! It is an inward and spiritual act!” is apt to appear as a solid moralist standing out against a generation of superficial theorizers: we see him as he sees himself, surveying the invisible depths of ethical space, with all the distinction of a specialist in the sui generis. Yet he provides Hippolytus with a let-out, the bigamist with an excuse for his “I do” and the welsher with a defense for his “I bet”. Accuracy and morality alike are on the side of the plain saying that our word is our bond. (Austin 1976, p.10)

No exame de noções como “promessa”, fica claro que várias condições podem não estar satisfeitas para que a promessa seja efetiva, menos o fato de que quem efetivamente prometeu não tivesse a “intenção” de assim fazê-lo. Este contexto mostra uma situação onde a minha descrição sobre “motivos” – “os quais apenas eu poderia descrever corretamente” não conta para nada – quando muito, constitui uma desculpa que dificilmente poderia ser trazida à baila com êxito para tentar explicar meus motivos para não ter cumprido o prometido.

Considerações finais

A estrutura do argumento final apresentado nesta dissertação é simples. Teorias perceptuais da introspecção dão suporte à visão não-relacional do argumento moral. Esta visão tem quatro características básicas.

A primeira destas características é a criação de um ponto de vista de primeira pessoa que isola o agente moral do contato com outros agentes e desconsidera a adoção da perspectiva de terceiros na atividade de avaliação moral. Somente para efeito de brevidade, chamaremos esta característica de perspectivismo isolacionista.

A segunda característica dá prioridade ao exame dos motivos ou intenções da ação em detrimento de uma avaliação sobre as consequências, previstas ou não, da ação moral. Em homenagem a sua filiação estoica, chamaremos esta característica de motivacionismo.

A terceira marca relevante da visão não-relacional é considerar a família de conceitos psicológicos utilizados na explicação motivacionista do argumento moral (“intenção”, “impulso”, “aversão”, “motivo”) como descrevendo estados mentais ocorrentes. Nomearemos esta característica como descritivismo.

A quarta e última característica favorece a ideia de que o sujeito moral tem um acesso privilegiado às suas fontes motivacionais onde apenas ele as poderia descrever corretamente. Daremos o nome de incorrigibilidade deliberativa a este aspecto.

Temos bons argumentos para rejeitar as teorias perceptuais da introspecção. Toda análise dos argumentos das *Royce Lectures* de Shoemaker teve o objetivo de fixar este ponto.

A visão não-relacional do argumento moral cai ou permanece em pé junto com as teorias perceptuais. Logo, temos que desconsiderar o

perspectivismo isolacionista, o motivacionismo, o descritivismo e a incorrigibilidade deliberativa como instrumentos adequados para explicar o funcionamento do argumento moral.

Seria uma tarefa interessante, mas que infelizmente não realizei aqui, identificar estas quatro características da visão não-relacional em teorias morais particulares, para além dos poucos autores que utilizamos para exemplificar os pontos em questão. Além desta exemplificação através de um exame da história da filosofia moral, é seguro afirmar que algumas destas características estão enraizadas em expressões que utilizamos cotidianamente. E arrisco a dizer que estas características estão presentes no próprio desenvolvimento, lento e gradual, da noção moderna de identidade do sujeito moral. Parte das reflexões do capítulo 4.1 foram consequências de uma leitura atenta de *Sources of the Self* de Charles Taylor (1989) onde temos um amplo panorama dos aspectos que moldaram esta identidade. Não podemos esquecer que a crítica ao modelo perceptual da introspecção é algo muito recente na escala dos acontecimentos filosóficos. E nada que tenha durado tanto tempo de forma hegemônica abandona a cena sem deixar marcas.

Um das perguntas centrais desta dissertação era sobre a relação entre teorias da introspecção e conceitos morais. A estrutura do argumento geral mostra que esta relação não só existe como é interna. Qualquer teoria sobre a moralidade pressuporá teses sobre atribuição e autoatribuição de atitudes proposicionais.

É claro que esta estrutura geral, assim descrita, não capta todos os pontos que considere importantes nesta discussão. Ela apenas funciona como um eixo que articula todos os temas aqui tratados. No tratamento dos textos de William James, em especial, toquei em várias outras questões que mereceriam uma atenção fora dos limites desta dissertação. Por mais que sintam-me obrigado a rejeitar a definição de introspecção que ele propõe, vários de seus temas são independentes da aceitação do modelo perceptual e procurei abordá-los, mesmo reconhecendo que apontariam para direções muito além dos nossos

interesses imediatos. A análise tripartite do *self* jamesoniano tem ressonâncias morais extremamente interessantes e que não cabem dentro da visão não-relacional. William James é um destes autores que carregamos para sempre e deixar promissórias em aberto na análise de seus textos é uma obrigação de honestidade intelectual.

O uso de autores da psicologia introspeccionista do final do séc. XIX, além de servir ao propósito de fixar teses do modelo perceptual, cumpre uma função específica dentro desta dissertação. Na literatura filosófica sobre o tema da introspecção, ou de modo mais geral, sobre autoconhecimento, é comum nos depararmos com estas expressões: “acesso privilegiado”, “autoridade da primeira pessoa”, “incorrigibilidade”, “infalibilidade”, “invulnerabilidade” e outras similares. Na maior parte das vezes não é uma tarefa simples definir o que exatamente os autores estão entendendo por uma ou mais destas expressões. Além disso, me parece claro, revisando a bibliografia sobre o tema, que estas teses girando em torno da infalibilidade da introspecção são ficções exclusivas da filosofia que, em nenhum momento foram supostas na Psicologia, nem mesmo no apogeu da metodologia introspeccionista. Abordar os textos de William James e James Sully cumpre o objetivo de mostrar que os primeiros teóricos do moderno desenvolvimento da Psicologia não compraram toda epistemologia “cartesiana” da introspecção.

Minha principal suspeita é de que a presunção de acesso privilegiado aos estados de consciência é uma projeção indevida de determinadas características gramaticais do discurso de primeira pessoa no processo introspectivo, em especial uma projeção descuidada das características da imunidade ao erro pela identificação equívoca (IEM) nos processos de autoatribuição de estados mentais.

Sydney Shoemaker defende uma posição próxima à nossa quando considera a plausibilidade intuitiva da tese da infalibilidade:

“To some extent, I think, the attractiveness of the infallibility claim is the product of the compelling force of Descartes' "Cogito argument." An

interesting slide occurs in the Second Meditation. Early on Descartes asserts the indubitability of "I think," along with "I exist." A few pages later he asserts, in effect, the indubitability of a wide range of first-person mental state ascriptions, including such judgments as that it seems to one that one sees light, that one hears noise and that one feels heat, this on the grounds that what these ascribe "is simply thinking." Recent philosophers who have reflected on why "I think" is immune even to the radical doubt of the First Meditation, based on the idea that one may be deceived by an all powerful demon, have generally seen that "It seems to me that I feel heat" cannot be immune to doubt for the same reason. "I think" is indubitable for a logical reason; it is a logically necessary condition of my being deceived about anything that I think, since being deceived is a matter of having false beliefs, which in turn is a special case of thinking - in the sense of "think" in question. "I think," like "I exist," is necessarily self-verifying, in the sense that it is a necessary condition of its being asserted, or even entertained in thought, that it be true. But it is obvious that attributions to oneself of particular thoughts, such as "I think that I am breathing," and also attributions to oneself of perceptual states, such as "It seems to me that I feel heat," do not have this status. The truth of these is not a necessary condition of their being asserted or entertained. Nevertheless, the slide from claiming the indubitability of "I think" to claiming the indubitability of propositions ascribing particular thoughts and conscious states is a very natural one." (Shoemaker 1996, *First-person access*, p. 52-53, meu sublinhado)

Nesta passagem Shoemaker fala de uma transição cartesiana na Segunda Meditação, passando da indubitabilidade do *Cogito, ergo sum* para a pretensa indubitabilidade das autoatribuições de estados mentais. Creio que Shoemaker têm bons argumentos para afirmar que as razões pelas quais o *Cogito* é indubitável são diferentes das razões para alegar indubitabilidade para proposições como "Penso que estou respirando" ou " Parece-me que sinto calor". Por outro lado, Tyler Burge - outro racionalista, de acordo com a nomenclatura que estamos usando nesta dissertação - mostrou o importante papel que este conjunto especial de proposições (*cogito-thoughts*) joga em nossa compreensão do autoconhecimento e da agência racional.⁵

⁵ Burge 1996 "Our Entitlement to Self-knowledge"

Não pretendo discutir estas teses aqui, mas antes sugerir outro tipo de transição, que ocorre já na Primeira Meditação cartesiana, deslocando o leitor para um cenário de isolamento. Dentre as características da visão não-relacional do argumento moral, creio que a mais fundamental seja o perspectivismo isolacionista. As conexões que traçamos na seção 4.1 entre esta característica e o surgimento da noção de consciência moral são especialmente importantes para a compreensão do cenário que estamos tentando descrever nesta dissertação.

Várias são as questões que permanecerão no horizonte de nossa reflexão sobre estes temas. Em especial, gostaríamos de lembrar ao nosso leitor da importância que exercícios confessionais em primeira pessoa desempenharam na história da filosofia. E lembrar, igualmente, que estes exercícios estão intimamente conectados com o perspectivismo isolacionista. Exemplos desta tendência não faltam, passando por Santo Agostinho, Rousseau, Thoreau e Kierkegaard – para ficarmos apenas com alguns dos autores mais representativos desta fixação no ponto de vista da primeira pessoa em questões de deliberação e avaliação moral.

Uma das questões que manteremos, por excelência, em nosso caminho, será a leitura atenta de um autor que tornou emblemática esta fixação moral no ponto de vista da primeira pessoa. Queremos dar relevo em nossos estudos de doutorado à analogia estrutural entre as *Meditações* de Descartes e o texto de *Walden*, de Henry David Thoreau, mantendo no pano de fundo a lembrança da filiação de Descartes à tradição dos exercícios espirituais tão como descritos por Pierre Hadot em *Exercices Spirituels et Philosophie Antique*

Por um lado, a operação da dúvida hiperbólica, por outro, o experimento de dois anos dois meses e dois dias de Thoreau vivendo às margens do lago Walden: exercícios que trazem a pretensão de estabelecer algo de firme, tanto no campo da certeza teórica, quanto no domínio da ação moral. Todavia, o início destas experiências pressupõem cenários temporários de atuação e esta maquinaria cenográfica tem algumas engrenagens em comum. A mais destacada delas, mas certamente não a única, é o afastamento do convívio regular com outras pessoas. Nossa companhia nestas jornadas será tão acidental

quanto os eventuais visitantes que passaram pela cabana de Thoreau durante seu experimento. Lembremos como Descartes inicia a Primeira Meditação:

Agora, pois, que meu espírito está livre de todos os cuidados, e que consegui um repouso assegurado numa pacífica solidão, aplicar-me-ei seriamente e com liberdade em destruir em geral todas as minhas antigas opiniões. (Descartes 1641/1973, p.115)

E agora Thoreau com a palavra, na primeira sentença de Walden:

When I wrote the following pages, or rather the bulk of them, I lived alone, in the woods, a mile from any neighbor, in a house which I had built myself, on the shore of Walden Pond, in Concord, Massachusetts, and earned my living by the labor of my hands only. I lived there two years and two months. At present I am a sojourner in civilized life again. (Thoreau 1854 /2004, p.1)

O que pode ensinar este retiro voluntário, este isolamento programado? Será esta uma condição necessária para a reflexão filosófica? O ponto central aqui é a relação entre o discurso filosófico, seja no domínio da certeza sobre os fundamentos do conhecimento, seja no campo da certeza sobre os fundamentos da ação moral, e os condicionantes da vida cotidiana expressos em nossos compromissos com o diário, com o senso comum. Depois de dar início à encenação do argumento filosófico, será que consigo voltar pelo mesmo trajeto, caminhando sem esforço nos rastros de minhas pegadas ou me será necessário uma longa jornada para trazer minhas palavras de volta para casa? Estas são apenas algumas das perguntas que esta analogia nos impõe. Na medida em que explorarmos as semelhanças, teremos outros interrogantes à espreita. De modo mais dramático, o tema da amizade aparecerá como problema fundamental não só para o pensamento de Thoreau, mas para a visão não-relacional do argumento moral como um todo. Não se trata apenas de refazer o caminho de volta, mas de saber se posso fazê-lo sozinho.

Referências

- Anscombe, Elizabeth. 1981. *The Collected Philosophical Papers of G.E.M. Anscombe Voll.III*, University of Minnesota Press.
- Armstrong, David. 1968. *A Materialist Theory of the Mind*. Routledge and Kegan Paul.
- Austin, J.L. 1970. *Philosophical Papers*. Oxford University Press.
- Austin, J. L. 1976. *How To Do Things With Words*. Oxford University Press.
- Baker, Gordon e Morris, Katherine. 1996. *Descartes' Dualism*. Routledge.
- Bar-On, Dorit. 2004. *Speaking My Mind – Expression and Self-Knowledge*. Oxford University Press.
- Bilgrami, Akeel. 2006. *Self-Knowledge and Resentment*. Harvard University Press.
- Boring, E. G. 1953. *A history of introspection*. Psychological Bulletin 50: 169-189.
- Brentano, Franz. 1874 / 1995. *Psychology from an Empirical Standpoint*. Routledge & Kegan Paul
- Brook, Andrew e De Vidi, Richard (Ed). 2001. *Self-Reference and Self-Awareness*. John Benjamins Publishing Company.
- Burge, Tyler. 1996. *Our Entitlement to Self-knowledge*. Proceedings of the Aristotelian Society 96: 91 – 116.
- Butler, Joseph. 1726/2006. *The Works of Bishop Butler*. University of Rochester Press.
- Cavell, Stanley. 2004. *Cities of Words*. Harvard University Press.
- Carruthers, Peter. 2011. *The Opacity of Mind*. Oxford University Press.

- Coope, Christopher, P. Geach, Timothy Potts e Roger White. 1970. *The Wittgenstein Workbook*. Basil Blackwell.
- Descartes, Rene.1641/1973 .*Obra Escolhida*. Difusão Européia do Livro.
- Edgley, Roy. 1969. *Reason in Theory and Practice*. Hutchinson.
- Evans, Gareth. 1982. *The Varieties of Reference*. Oxford University Press.
- Frankish, Keith. 2004. *Mind and Supermind*. Cambridge University Press.
- Gallois, André. 1996. *The World without, the mind within: an essay on first-person authority*. Cambridge University Press.
- Gertler, Brie. 2011. *Self-knowledge*. Routledge.
- Goldman, Alvin. 2006. *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. Oxford University Press.
- Goodman, Russell B. 2004. *Wittgenstein and William James*. Cambridge University Press.
- Hacker, P.M.S. 1997. *Wittgenstein on Human Nature*. Phoenix.
- Hadot, Pierre .1981. *Exercices Spirituels et Philosophie Antique*. Études Augustiniennes.
- Hicks, G. Daves. 1927. Proceedings of the Aristotelian Society, Supplementary Volumes, Vol. 7: 55-97
- Hintikka, Jaakko. 1962. *Knowledge and Belief*. Cornell Universtity Press.
- James, William. 1890 / 1983. *The Principles of Psychology*. Harvard University Press.
- James, William. 1897. *The Will to Believe and Other Essays in Popular Philosophy*. The Project Gutenberg Ebook.
- James, William. 1902/2002. *The Varieties of Religious Experience*. Routledge.

- Langston, Douglas. 2001. *Conscience and other Virtues*. The Pennsylvania State University Press.
- Locke, John. (1689/1975) *An Essay Concerning Human Understanding*. Oxford University Press.
- Lycan, William. 1996. *Consciousness and Experience*. MIT Press.
- Lyons, William. 1986. *The Disappearance of Introspection*. MIT Press.
- Malcolm, Norman. 1971. *Problems of Mind*. George Allen & Unwin.
- McLaughlin, B; Rorty, A (Ed). 1988. *Perspectives on Self-Deception*. University of California Press.
- Nichols, Shaun e Stich, Stephen. 2003. *Mindreading – An Integrated Account of Pretence, Self-Awareness and Understanding Other Minds*. Oxford University Press.
- Prinz, Jesse. 2007. *The Emotional Construction of Morals*. Oxford University Press.
- Putnan, Ruth Anna (Ed). 1997. *The Cambridge Companion to William James*. Cambridge University Press.
- Russell, Bertrand. 1921. *The Analysis of Mind*. George Allen & Unwin Ltd.
- Ryle, Gilbert. 1949. *The Concept of Mind*. Barnes and Nobles.
- Schwitzgebel, Eric. 2011. *Perplexities of Consciousness*. MIT Press.
- Seddon, Keith. 2005. *Epictetu's Handbook and the Tablet of Cebes – Guides to Stoic Living*. Routledge.
- Shoemaker, Sydney. 1996. *The First-person perspective and other essays*. Cambridge University Press.
- Smithies, Declan; Stoljar, Daniel (Ed). 2012. *Introspection and Consciousness*. Oxford University Press.

Sully, James. 1881 /1887. *Illusions – A Psychological Study*. Kegan Paul, Trench & Co.

Strawson, Galen. 2009. *Selves–An Essay in Revisionary Metaphysics*. Oxford University Press.

Taylor, Charles. 1989. *Sources of the Self*. Harvard University Press.

Thoreau, Henry D. 1854/2004. *Walden – A fully annotated edition*. Yale University Press.

Titchener, E.B. 1912. Prolegomena to a study of introspection. *American Journal of Psychology* 23: 485-508.

Wittgenstein, Ludwig.1953 / 2009. *Philosophical Investigations*. Blackwell Publishing Ltd.

Wittgenstein, Ludwig. 1958. *The Blue and Brown Books*. Blackwell Publishers.