

Introspecção

Eric Schwitzgebel



INTROSPECÇÃO

ERIC SCHWITZGEBEL

Tradução de Cíntia Fernandes Marcellos, Marcus
Vinícius do Amaral Gama Santos, Kamilly Gomes da
Silva, Giulia Nogueira Franca, Gunther Mafra
Guimarães e Yuri Pereira Antunes Vieira

INTROSPECÇÃO

Editora do Portal História da Psicologia

Rio das Ostras, RJ

2025

EDITORA DO PORTAL HISTÓRIA DA PSICOLOGIA

A editora do Portal História da Psicologia é parte integrante do Portal História da Psicologia

EDITOR

André Elias Morelli Ribeiro

CONSELHO EDITORIAL

André Elias Morelli Ribeiro

Arthur Arruda Leal Ferreira

Luiz Eduardo Prado da Fonseca

Marcus Vinícius do Amaral Gama Santos

© 2025

É proibida a reprodução total ou parcial desta obra sem autorização expressa da Editora do Portal História da Psicologia

© Stanford Encyclopedia of Philosophy, 2019

Publicado em acordo com a Universidade de Stanford

Título original: *Introspection*

Equipe de realização

Editor Responsável: André Elias Morelli Ribeiro

Revisão final: André Elias Morelli Ribeiro

Capa: André Elias Morelli Ribeiro, com imagens geradas por inteligência artificial (Midjourney AI)

Projeto gráfico e diagramação: André Elias Morelli Ribeiro

Ficha catalográfica elaborada pela bibliotecária Pâmella Priscilla Negrão
Braga CRB-7/6062

S415i Schwitzgebel, Eric
 Introspecção. / Eric Schwitzgebel; [coordenação André Elias Morelli Ribeiro]; tradução de Cíntia Fernandes Marcellos, Marcus Vinícius do Amaral Gama Santos, Kamilly Gomes da Silva, Giulia Nogueira Franca, Gunther Mafrá Guimarães, Yuri Pereira Antunes Vieira. – Rio das Ostras, RJ: Editora do Portal História da Psicologia, 2025.
 [xiv], 165p.

E-book.

ISBN: 978-65-997325-5-3

DOI: 10.5281/zenodo.14937499

Bibliografia p. 133-165

1. Teoria do conhecimento. 2. Epistemologia. 3. Filosofia.
I. Ribeiro, André Elias Morelli (Coordenador). II. Marcellos, Cíntia Fernandes (Tradutor). III. Santos, Marcus Vinícius do Amaral Gama (Tradutor). IV. Silva, Kamilly Gomes da (Tradutor). V. Franca, Giulia Nogueira (Tradutor). VI. Guimarães, Gunther Mafrá (Tradutor). VII. Vieira, Yuri Pereira Antunes (Tradutor). VIII. Título.

CDD 121

Este livro é a tradução do verbete *Introspection*, publicado na *Stanford Encyclopedia of Philosophy*

A presente versão foi publicada em 2 de fevereiro de 2010 e revisada em 18 de outubro de 2019

A versão arquivada deste verbete pode ser encontrada em:

<https://plato.stanford.edu/archives/win2019/entries/introspection>

Nota do editor

O processo de tradução e construção deste livro foi relativamente longo e difícil. Inicialmente, a ideia de trabalhar com o texto de Schwitzgebel se fundamentava numa avaliação parcial e bastante equivocada da complexidade do material, a ponto de ser entregue para duas alunas calouras, graduandas em psicologia, Giulia Nogueira Franca e Kamilly Gomes da Silva, visando atribuição de notas para um componente curricular. A confiança na habilidade de ambas em efetivar a tarefa vinha de seus contatos prévios com o idioma inglês.

Rapidamente, ambas procuraram o docente responsável, queixando-se da sua dificuldade para realizar a tradução a contento, considerada por demais desafiadora. Ao analisar mais detidamente o material, a conclusão foi inevitável: era necessário o emprego de outro conjunto de tradutores para lidar com a delicada tradução. Decidiu-se pela continuidade de ambas na execução da tarefa, em respeito àquilo que já havia sido feito, mas desta vez indicando todos os problemas e dificuldades, que seriam debatidas e resolvidas por outro comitê. Isto tudo se desenrolou no ano de 2022.

Seguiu-se então a convocação de um pequeno, mas engajado grupo de pesquisadores em história da psicologia,

bastante habituados com o debate acadêmico em inglês, composto por Marcus Vinícius do Amaral Gama Santos, Gunther Mafra Guimarães, Yuri Pereira Antunes Vieira e André Elias Morelli Ribeiro, com a colaboração eventual de Luiz Eduardo Prado da Fonseca. Já na apresentação do material, Santos - então estudante de filosofia e mestrando em psicologia - indicou algumas questões relacionadas a termos, conceitos e expressões filosóficas que exigiriam conhecimentos mais amplos do assunto. Ato contínuo, o mesmo pesquisador apontava sutilezas nas frases do autor, que se relacionavam com o debate proposto e que exigiam uma atenção maior e até mesmo certa criatividade na tradução. As características do material se mostravam crescentemente desafiadores, bem como o entendimento da sua relevância para o público falante de português.

Este grupo viu-se então incapaz de resolver os problemas de tradução que se apresentavam. Conforme as discussões se desenrolavam, aos itens indicados pelas primeiras tradutoras como dificuldades para a tradução, adicionou-se muitos outros, o que mostrou que as estratégias adotadas para resolver os problemas de tradução ainda não eram eficientes o suficiente.

A solução foi a convocação de pesquisadores especialistas no tema da introspecção. Alguns foram convidados e, após avaliar

a dificuldade e complexidade da tarefa, vários recusaram. Até que um destes convidados sugeriu o nome da professora Dra. Cíntia Fernandes Marcellos, do Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais. Tendo estudado as ideias de Edward Titchener em seu doutorado, ela havia traduzido, junto do saudoso professor Dr. Saulo de Freitas Araújo, textos do autor, publicados sob o título “Edward B. Titchener, Psicologia, ciência e tecnologia”, lançado pela editora Hogrefe. Trata-se de um imenso desafio, que ela e Araújo realizaram com exemplar habilidade. O simples fato de a Dra. Marcellos ter trabalhado com o Dr. Araújo numa tradução tão delicada já servia de indicador de sua capacidade.

Quando contatada, ela aceitou o desafio, tomou a tradução em suas mãos e realizou uma série de melhorias, modificações, adaptações, tornando o texto realmente viável para leitores de língua portuguesa. Além disso, adicionou notas de discussão sobre detalhes mais delicados na tradução que recebeu. Receosa com a precisão de alguns dos pontos destacados, ela informa o editor a necessidade de análise. É quando Marcus Vinícius do Amaral Gama Santos, que havia participado do segundo grupo de tradução, retorna à cena e propõe uma série de discussões e sugestões sobre os destaques de Cíntia. Em alguns deles, não

houve consenso entre ambos, em outros, a discussão entre os colaboradores foi suficiente, sendo empregada nesta versão final.

Nas questões sem consenso, contudo, coube ao editor resolver as querelas. Este entende que não é de bom tom abusar do recurso de explicações sobre a tradução ao longo do texto, em geral bastando ao apreciador do prato crer na lisura e capacidade da cozinheira. Desta feita, para não tornar o texto carregado por considerações de tradução, as sintetizo aqui visando o bem-estar do leitor.

Em primeiro lugar, a palavra introspecção ganhou uma variedade de flexões verbais, tornando-se, quando necessário, “introspectar” ou “introspeccionar” etc., o que confere ao texto uma familiaridade útil para o leitor lusófono.

O conceito de *self-intimation*, que até nomeia uma das seções do verbete, apresentou sérias dificuldades para os tradutores. Foram encontradas algumas formas diferentes de traduzir a expressão na literatura, e a escolhida para o presente volume foi *autorrevelação*. Já o conceito de *self-fulfilling* também foi objeto de discussão pelos tradutores, que optaram pela expressão *autorrealizadora*. Outros problemas e dificuldades na tradução emergem no texto acompanhados do original entre colchetes, para que o leitor compreenda melhor as escolhas e as intenções originais do autor.

Por fim, mas não menos importante, os editores e tradutores dedicam o trabalho empreendido neste volume ao Prof. Dr. Saulo de Freitas Araújo. Saulo, como seus colegas e amigos costumavam chamá-lo, foi um grande exemplo de dedicação e qualidade na pesquisa histórica da psicologia e da filosofia da psicologia. Reconhecido internacionalmente como um dos maiores especialistas em Wundt – um feito extraordinário para um falante nativo de português – escreveu de forma brilhante sobre Titchener, Fechner, William James, Christian Wolff, John Watson, historiografia da psicologia, filosofia da história, entre outros. Seus textos eram rigorosos e aguardados com ansiedade por todos, sempre discutidos e ocupando um lugar especial no debate da comunidade de historiadores da psicologia.

Em específico, sua figura nos inspira a executar a presente tradução. Ele coordenou uma coleção de traduções de clássicos da história da psicologia e traduziu ele próprio muitos materiais importantes, defendendo a tradução como um dever acadêmico e uma necessidade de todos os envolvidos no campo. Traduzir é derrubar barreiras e permitir novos conhecimentos e conexões, tarefas empreendidas majestosamente pelo Saulo, que nos inspira a continuar nesta tarefa.

Sumário

1. Características Gerais da Introspecção....17
 - 1.1 Características necessárias para um processo introspectivo....17
 - 1.2 Os alvos da introspecção....25
 - 1.3 Os produtos da introspecção....28
2. Abordagens Introspectivas Versus Não-Introspectivas do Autoconhecimento....29
 - 2.1 Abordagem de paridade eu/outro....29
 - 2.1.1 Abordagens da observação comportamental....30
 - 2.1.2 Abordagens da teoria da teoria....32
 - 2.1.3 Restrições à paridade....34
 - 2.2 Abordagens de autodeteção....37
 - 2.2.1 Abordagens de monitoramento simples....41
 - 2.2.2 Abordagens de monitoramento multiprocesso....45
 - 2.3 Introspecção sem autodeteção?...49
 - 2.3.1 Autorrealização e contenção....52
 - 2.3.2 Automodelagem....57
 - 2.3.3 Expressivismo....61
 - 2.3.4 Transparência....63

2.4	Pluralismo introspectivo....	73
3.	O Papel da Introspecção na Psicologia Científica....	76
3.1	A ascensão da psicologia introspectiva como ciência....	76
3.2	O ceticismo inicial sobre a observação introspectiva....	79
3.3	O declínio da introspecção científica....	83
3.4	A reemergência da introspecção científica....	86
4.	A Precisão da Introspecção....	92
4.1	Variedades de privilégio....	92
4.1.1	Variedades de perfeição: infalibilidade, indubitabilidade, incorrigibilidade e autorrevelação....	93
4.1.2	Garantias mais fracas....	100
4.1.3	Privilégio sem garantia....	106
4.2	Evidência empírica sobre a precisão da introspecção....	109
4.2.1	Das causas das atitudes e comportamentos....	110
4.2.2	De atitudes....	117
4.2.3	Da experiência consciente....	126
	Referências Bibliográficas....	133

INTROSPECÇÃO, no contexto da filosofia contemporânea da mente, refere-se ao processo pelo qual adquirimos conhecimento sobre nossos próprios estados ou processos mentais, sejam eles atuais ou ocorridos muito recentemente. É claro que você pode aprender sobre sua mente da mesma forma que você aprende sobre a mente alheia - por meio da leitura de textos de psicologia, da observação de expressões faciais (em um espelho), do exame de leituras de atividade cerebral, da observação de comportamentos passados - mas geralmente pensa-se que você também pode conhecer sua mente *introspectivamente*, de um jeito que mais ninguém pode. Mas o que é exatamente a introspecção? Nenhuma definição simples é amplamente aceita.

A introspecção é um conceito-chave na epistemologia, uma vez que o conhecimento introspectivo é frequentemente compreendido como particularmente seguro, talvez até imune à dúvida cética. O conhecimento introspectivo também é frequentemente considerado mais imediato ou direto do que o conhecimento sensorial. Estas duas supostas características da introspecção têm sido citadas em apoio à ideia de que o conhecimento introspectivo pode servir como uma base ou fundamento para outros tipos de conhecimento.

A introspecção também é central para a filosofia da mente, tanto como um processo digno de estudo por si só como também uma corte de apelação para outras reivindicações sobre a mente. Filósofos da mente oferecem uma variedade de teorias sobre a natureza da introspecção; e reivindicações filosóficas sobre consciência, emoção, livre arbítrio, identidade pessoal, pensamento, crença, representação [*imagery*], percepção, e outros fenômenos mentais são frequentemente considerados como tendo consequências introspectivas ou serem suscetíveis à verificação introspectiva. Por razões similares, psicólogos empíricos também discutiram a precisão dos juízos introspectivos e o papel da introspecção na ciência da mente.

1. Características Gerais da Introspecção

1.1 Características necessárias para um processo introspectivo

A introspecção é comumente considerada como um processo por meio do qual aprendemos sobre nossos próprios estados ou processos mentais em curso ou muito recentes. No entanto, nem todos esses processos são introspectivos: alguns diriam que você realizou a introspecção se você se der conta que está com raiva ao ver sua expressão facial no espelho. Contudo, é incerto e controverso saber o que mais é necessário para um processo ser qualificado como introspectivo. Uma abordagem relativamente restrita da introspecção talvez requeira o envolvimento da atenção ao estado mental em curso e sua detecção direta; mas muitos filósofos pensam que uma atenção aos estados mentais ou sua detecção direta é impossível ou pelo menos não está presente em muitas instâncias paradigmáticas da introspecção.

Para um processo ser qualificado como “introspectivo”, como o termo é comumente usado na filosofia da mente

contemporânea, ele deve atender minimamente as três seguintes condições:

1. *A condição de mentalidade*: a introspecção é um processo que gera, ou visa gerar, conhecimento, juízos, ou crenças sobre eventos, estados ou processos *mentais*, e não sobre assuntos exteriores a uma mente, pelo menos não diretamente. A respeito disso, é diferente dos processos sensoriais que normalmente fornecem informações sobre eventos externos ou sobre aspectos não mentais do corpo do indivíduo. O limite entre o conhecimento introspectivo e não introspectivo pode começar a se confundir em relação ao autoconhecimento corporal, tal como o conhecimento proprioceptivo sobre a posição dos membros, ou o conhecimento nociceptivo sobre as dores. Mas, em princípio, a parte introspectiva de tais processos, referentes a julgamentos sobre a mente - por exemplo, que se tem a sensação como se os braços foram cruzados ou de dor localizada nos dedos do pé - podem ser distinguidos dos julgamentos não introspectivos de que os braços estão, de fato, cruzados ou que o dedo do pé está sendo beliscado.
2. *A condição de primeira pessoa*: a introspecção é um processo que gera, ou visa gerar, conhecimento,

juízos ou crenças sobre a *própria mente* e a de mais ninguém, pelo menos não diretamente. Qualquer processo que, de maneira semelhante, gere conhecimento sobre a própria mente e a de outros não é, por essa lógica, um processo introspectivo. (Alguns filósofos contemplaram casos peculiares ou de ficção científica em que podemos fazer introspecção diretamente ao conteúdo da mente de outras pessoas - por exemplo, por telepatia ou quando os cérebros de duas pessoas estão diretamente conectados - mas a devida interpretação desses casos é discutível. Veja, por exemplo, Gertler 2000).

3. *A condição de proximidade temporal*: a introspecção é um processo que gera conhecimento, crenças ou juízos apenas sobre a vida mental *atualmente em curso*; ou, alternativamente (ou talvez além disso) sobre a vida mental *imediatamente passada* (ou mesmo futura), dentro de uma certa janela estreita de tempo (às vezes chamada de presente ilusório). Você pode saber que ontem, durante sua caminhada matinal, estava pensando em Montaigne, mas não pode saber desse fato apenas pela introspecção atual - embora talvez possa saber introspectivamente que atualmente tem uma lembrança

vívida de ter pensado em Montaigne. Da mesma forma, você não pode saber apenas por introspecção que se sentirá deprimido se o seu candidato favorito perder as eleições em novembro - embora talvez você possa saber introspectivamente qual sua atitude atual em relação à eleição, ou que emoção começa a surgir em você quando considera os possíveis resultados. Se o alvo da introspecção é melhor pensado como a vida mental atual, ou a vida mental imediatamente passada, pode depender do modelo de introspecção: em modelos de autodeteccção de introspecção, segundo o qual a introspecção é um processo causal que envolve a detecção de um estado mental (ver a seção 2.2, abaixo), é natural supor que um breve lapso de tempo transcorrerá entre a ocorrência do estado mental, que é o alvo introspectivo, e o juízo introspectivo final sobre esse estado, o que “convoca” (mas não implica estritamente) a ideia de que os juízos introspectivos geralmente dizem respeito a estados imediatamente passados. Sobre modelos de automodelagem e autosatisfação de introspecção, segundo os quais os juízos introspectivos criam ou incorporam o próprio estado introspectivo (ver seções 2.3.1 e 2.3.2, abaixo), parece mais natural pensar que o

objetivo da introspecção é a vida mental atual de alguém ou, talvez, até mesmo de um futuro próximo.

Poucos filósofos contemporâneos da mente chamariam um processo de “introspectivo” se ele não atendesse alguma versão das três condições acima, embora na linguagem comum a condição de proximidade temporal possa, às vezes, ser violada. (por exemplo, no discurso comum, podemos descrever como “introspectivo” um processo de pensamento sobre o porquê você desistiu de um relacionamento mês passado ou se você é realmente gentil com seus filhos como pensa que é.) Porém, muitos filósofos da mente vão resistir a chamar um processo que atenda essas três condições de “introspecção”, a menos que também satisfaça algumas, ou todas, das três condições a seguir:

4. *A condição de diretividade*: a introspecção produz juízos ou conhecimentos sobre os próprios processos mentais atuais de forma relativamente *direta* ou *imediate*. É difícil articular exatamente o que a diretividade ou o imediatismo implica no contexto atual, mas alguns exemplos devem tornar relativamente clara a importância dessa condição. Reunir informações sensoriais sobre o mundo e, em seguida, tirar conclusões teóricas com base

nessas informações não deve, de acordo com essa condição, ser considerado introspectivo, mesmo que o processo atenda às três condições acima. Ver que um carro está a seis metros a sua frente e então deduzir, a partir desse fato sobre o mundo externo, que você está tendo uma experiência visual de certo tipo, não se caracteriza, por essa condição, como introspectivo. No entanto, como veremos na seção 2.3.4, abaixo, aqueles que adotam as teorias de transparência da introspecção podem rejeitar, pelo menos, fortes formulações dessa condição.

5. *A condição de detecção*: a introspecção envolve algum tipo de *sintonização a* ou *detecção de* um estado ou evento mental *pré-existente*, no qual o juízo ou conhecimento introspectivo é (quando tudo vai bem) *causalmente*, mas não *ontologicamente*, dependente do estado mental alvo. Por exemplo, um processo que envolvesse criar o estado da mente que se atribui a si mesmo não seria introspectivo, de acordo com essa condição. Suponha que eu diga a mim mesmo em um discurso interior silencioso: “Estou dizendo a mim mesmo, em um discurso interior silencioso, ‘*haecceities of applesauce*’” sem nenhuma ideia prévia de como

pretendo completar a citação introduzida. Agora, o que eu digo pode ser verdade, e eu posso saber que é verdade, e eu posso conhecer sua verdade (em algum sentido) diretamente, por um meio pelo qual eu não poderia conhecer a verdade da mente de nenhuma outra pessoa. Ou seja, pode atender a todas as quatro condições acima e, no entanto, podemos resistir a chamar essa autoatribuição de introspectiva. As abordagens de autoconhecimento dos tipos automodelagem (Seção 2.3.2, abaixo), expressivismo (Seção 2.3.3, abaixo) e transparência (Seção 2.3.4, abaixo) enfatizam o quanto nosso autoconhecimento, em muitos casos, não envolve a detecção de estados mentais pré-existentes; e porque algo como a condição de detecção é aceito implicitamente ou explicitamente por muitos filósofos, alguns filósofos (incluindo alguns, mas não todos, que apoiam a ideia da automodelagem, expressivismo e/ou transparência) poderiam considerar como inapropriado pensar em tais abordagens do autoconhecimento como abordagens propriamente ditas da *introspecção*.

6. *A condição de esforço*: a introspecção não é *constante*, *sem esforço* e *automática*. Nós não estamos a cada minuto do dia realizando a introspecção. A introspecção envolve

algum tipo de reflexão especial sobre a própria vida mental que difere do fluxo comum de pensamentos e ações não autoreflexivas. A mente pode monitorar a si mesma regularmente e constantemente sem exigir nenhum ato especial de reflexão por parte do pensador - por exemplo, em um nível não-consciente, certas partes do cérebro e outros sistemas funcionais podem monitorar os acontecimentos de outras partes do cérebro e outros sistemas funcionais, e esse monitoramento pode atender a todas as cinco condições acima - mas esse tipo de coisa não é o que os filósofos geralmente têm em mente quando falam de introspecção. No entanto, essa condição, assim como as condições de diretividade e detecção, não é universalmente aceita. Por exemplo, filósofos que pensam que a experiência consciente requer algum tipo de monitoramento introspectivo da mente e que pensam na experiência consciente como uma característica mais ou menos constante de nossas vidas podem rejeitar a condição de esforço (Armstrong 1968,1999; Lycan 1996).

Embora nem todas as abordagens filosóficas apresentadas por seus autores como abordagens de “introspecção” atendam a todas as condições de 4 a 6, a maioria

atende a pelo menos duas delas. Devido às diferenças na importância atribuída às condições 4-6, não é incomum que autores com outras abordagens semelhantes de autoconhecimento difiram em sua vontade de descrevê-las como abordagens de “introspecção”.

1.2 Os alvos da introspecção

Abordagens de introspecção diferem no que eles tratam como os *alvos* apropriados do processo introspectivo. Nenhum grande filósofo contemporâneo acredita que toda a mentalidade está disponível para ser descoberta pela introspecção. Por exemplo, os processos cognitivos envolvidos no processamento visual primário e na detecção de fonemas são frequentemente considerados introspectivamente impenetráveis e, não obstante (em algum sentido importante), mentais (Marr 1983; Fodor 1983). Muitos filósofos também aceitam a existência de crenças ou desejos inconscientes, aproximadamente no sentido freudiano, que não estão introspectivamente disponíveis (por exemplo, Gardner 1993; Velleman 2000; Moran 2001; Wollheim 2003; a despeito disso veja Lear 1998). Embora no uso comum do inglês às vezes dizemos que estamos “introspectando” quando refletimos sobre nossos traços de caráter, os filósofos

contemporâneos da mente geralmente não acreditam que possamos introspectar diretamente traços de caráter da mesma forma que podemos introspectar alguns de nossos outros estados mentais (especialmente à luz de pesquisas sugerindo que às vezes temos pouco conhecimento de nossos traços, analisados em Taylor e Brown 1988; Paulhus e John 1998; Vazire 2010).

Os dois tipos de estados mentais passíveis de introspecção mais comumente citados são *atitudes*, tais como crenças, desejos, avaliações e intenções, e *experiências conscientes*, tais como as emoções, imagens e experiências sensoriais. (Esses dois grupos podem não ser total ou mesmo parcialmente, dissociados: dependendo de outros aspectos de sua visão, uma filósofa pode considerar algumas ou todas as experiências conscientes envolvendo atitudes e/ou ela pode considerar atitudes como coisas que são ou podem ser experimentadas conscientemente). Do fato (se for um fato) de que algumas atitudes são introspectáveis, não se segue, é claro, que todas as atitudes o são, ou do fato de que algumas experiências conscientes são introspectáveis, que todas as experiências conscientes o são. Algumas abordagens de introspecção focam em atitudes (por exemplo, Nichols e Stich 2003), enquanto outras focam em experiências conscientes (por exemplo, Hill 1991; Goldman 2006; Schwitzgebel 2012); e às vezes não é claro até que ponto os

filósofos anseiam que suas observações sobre a introspecção de um tipo de alvo se apliquem a outro tipo. Não há garantia de que o mesmo mecanismo ou processo esteja envolvido na introspecção de todos os diferentes alvos em potencial.

Genericamente, este artigo descreverá os alvos da introspecção como *estados mentais*, embora em alguns casos possa ser mais adequado pensar os alvos como processos em vez de estados mentais. Além disso, ao falar dos alvos da introspecção como *alvos*, nenhuma pressuposição se destina a uma visão de autodeteção da introspecção em oposição a uma visão de automodelagem ou de contenção ou expressivista (Veja a seção 2, abaixo). Os alvos são simplesmente os estados autoatribuídos como uma consequência do processo introspectivo se o processo funcionar corretamente, ou, se o processo introspectivo falhar, os estados que teriam sido autoatribuídos.

1.3 Os produtos da introspecção

Embora filósofos não tenham explorado a questão minuciosamente, as abordagens também diferem em relação aos *produtos* da introspecção. A maioria dos filósofos defende que a introspecção produz algo como crenças ou juízos sobre a própria mente, mas outros preferem caracterizar os produtos da

introspecção como “pensamentos”, “representações”, “consciência” ou algo semelhante. Para facilitar a exposição, esse artigo descreverá os produtos do processo introspectivo como juízos, sem pretender tomar como ponto pacífico a contraposição com visões conflitantes.

2. Abordagens Introspectivas Versus Não-Introspectivas do Autoconhecimento

Esta seção delineará várias abordagens do autoconhecimento. Nem todas merecem ser chamadas de introspectivas, mas uma compreensão da introspecção requer uma apreciação dessa diversidade de abordagens - algumas por causa do contraste que fornecem à introspecção propriamente dita e outras porque é discutível se deveriam ser classificadas como introspectivas. Essas abordagens não são exclusivas. Certamente há mais de um processo por meio do qual podemos obter autoconhecimento. Inevitavelmente, parte do mesmo território aqui abordado também é apresentado, de maneira bastante diferente, no verbete sobre autoconhecimento¹.

2.1 Abordagem de paridade eu/outro

As abordagens simétricas ou de paridade eu/outro do autoconhecimento tratam os processos pelos quais adquirimos conhecimento da nossa própria mente como essencialmente os mesmos que tratam os processos pelos quais adquirimos

¹ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/self-knowledge/>>

conhecimento das mentes de outras pessoas. Uma versão simplista dessa visão é que conhecemos tanto nossa própria mente quanto a mente de outros apenas observando o comportamento externo. Em tal visão, a introspecção, estritamente falando, é impossível, uma vez que a condição de primeira pessoa da introspecção (condição 2 na Seção 1.1) não pode ser atendida: não há processos distintos que gerem conhecimento apenas da própria mente. Os princípios behavioristas do século XX tendiam a encorajar essa visão, mas nenhum tratamento proeminente do autoconhecimento aceita essa visão em sua forma mais extrema e simples. Os defensores das abordagens de paridade às vezes caracterizam o conhecimento que temos de nossa própria mente como oriundo de “teorias” que aplicamos igualmente a nós mesmos e aos outros (como em Nisbett e Ross 1980; Gopnik 1993a, 1993b). Consequentemente, essa abordagem do autoconhecimento às vezes é chamada de *teoria da teoria*.

2.1.1 Abordagens da observação comportamental

Dentre os principais pesquisadores, a visão de Bem (1972), embora agora datada, talvez seja a que mais se aproxima de uma visão simples da paridade eu/outro, argumentando com base em pesquisas psicológicas que nosso conhecimento dos “estados

internos”, tanto do eu quanto dos outros, deriva em grande parte dos mesmos tipos de evidência comportamental e emprega os mesmos princípios de inferência. Nós notamos como nos comportamos e então inferimos as atitudes evidenciadas por esses comportamentos - e o fazemos mesmo quando realmente nos falta a atitude imputada. Bem cita uma pesquisa clássica em psicologia social sugerindo que, quando induzidas a realizar uma ação por uma pequena recompensa, as pessoas atribuirão a si mesmas uma atitude mais positiva em relação a essa ação do que quando são induzidas por uma grande recompensa (Festinger e Carlsmith 1959; veja também a Seção 4.2.2, abaixo). Quando percebemos que estamos fazendo algo com compensação mínima, inferimos uma atitude positiva em relação a essa atividade, assim como faríamos se vissemos outra pessoa realizando a mesma atividade com compensação mínima. De forma semelhante, talvez saibamos que gostamos de comida tailandesa porque notamos que às vezes atravessamos a cidade de carro para obtê-la; podemos saber que estamos felizes porque nos vemos ou nos sentimos sorrindo. Bem argumentou que a psicologia social falhou consistentemente em mostrar que temos qualquer acesso significativo a informações privadas que possam ir contra tais autoatribuições orientadas para o exterior. Na visão de Bem, se somos melhores em discernir nossas próprias

motivações e atitudes, é primariamente porque observamos mais o nosso próprio comportamento do que o de qualquer outra pessoa.

2.1.2 Abordagens da teoria da teoria

Nisbett, Wilson e seus coautores (Nisbett e Bellows 1977; Nisbett e Wilson 1977; Nisbett e Ross 1980; Wilson 2002) argumentam de forma semelhante pela paridade eu/outro em nosso conhecimento sobre as bases ou causas das nossas próprias atitudes e comportamentos e as dos outros, descrevendo casos em que as pessoas parecem apresentar pouco conhecimento dessas bases ou causas. Por exemplo, pessoas questionadas em um centro comercial suburbano sobre porque escolheram um determinado par de meias pareciam ignorantes da influência da posição sobre aquela escolha, inclusive negaram explicitamente aquela influência quando lhes foi sugerida. Pessoas solicitadas para avaliar várias características de supostos candidatos a emprego não sabiam que seus juízos sobre a flexibilidade do candidato foram muito influenciados por terem sido informadas de que o candidato havia derramado café durante a entrevista de emprego (consulte também a Seção 4.2.2, abaixo). Em tais casos, Nisbett e seus co-investigadores descobriram que as descrições dos participantes sobre as influências causais em seu próprio

comportamento espelhavam de perto as influências hipotetizadas por observadores externos. A partir dessa descoberta, eles inferem que o mesmo comanda as atribuições de primeira e terceira pessoa, um mecanismo que não envolve nenhum acesso privado especial às verdadeiras causas das atitudes e comportamentos de cada um e, em vez disso, se baseia fortemente em teorias psicológicas intuitivas.

Gopnik (1993a, 1993b; Gopnik e Meltzoff 1994) utiliza evidências psicológicas de desenvolvimento para apoiar uma teoria da paridade do autoconhecimento. Ela aponta para a evidência de que para uma ampla variedade de estados mentais, incluindo acreditar, desejar e fingir, as crianças desenvolvem a capacidade de atribuir esses estados a si mesmas na mesma idade em que desenvolvem a capacidade de atribuir esses estados a outros. Por exemplo, as crianças não parecem ser capazes de atribuir a si mesmas falsas crenças passadas (depois de terem sido enganadas pelo experimentador) antes do que podem atribuir falsas crenças a outras pessoas. Isso parece ser dessa forma mesmo quando essa falsa crença está em um passado muito recente, tendo acabado de ser revelada como falsa. De acordo com Gopnik, esse paralelismo pervasivo mostra que não nos é dado acesso introspectivo direto a nossas crenças, desejos, pretensões e similares. Em vez disso, devemos desenvolver uma

“teoria da mente”, à luz da qual interpretamos as evidências que garantem nossas autoatribuições. O surgimento de uma evidência imediata dos próprios estados mentais é, sugere Gopnik, apenas uma “ilusão de expertise: os experts se envolvem em todo tipo de teorização tácita que não reconhecem como tal – o jogador de xadrez expert para quem a força de uma jogada parece simplesmente dada visualmente, o médico que imediatamente intui câncer em um paciente. Uma vez que somos todos experts na atribuição de estado mental, não reconhecemos as camadas de teoria que sustentam o processo.

2.1.3 Restrições à paridade

A evidência empírica por trás das visões de paridade eu/outro permanece controversa (White 1988; Nichols e Stich 2003; Carruthers 2011). Ademais, embora Bem, Nisbett, Wilson e Gopnik salientem o paralelismo entre a atribuição do estado mental a si mesmo e aos outros e a natureza inferencial e teórica de tais atribuições, todos eles também deixam margem para um tipo de autoconsciência de natureza diferente da consciência que se tem da vida mental dos outros. Assim, nenhum endossa uma visão *puramente* simétrica ou de paridade eu/outro. Bem reconhece que o paralelismo só se mantém “na medida em que pistas internas são fracas, ambíguas ou não interpretáveis” (1972,

5). Com essa ressalva em mente, ele afirma que nosso autoconhecimento é “parcialmente” baseado em pistas externas. Nisbett e Wilson enfatizam que nos falta acesso apenas aos “processos” ou causas *subjacentes* ao nosso comportamento e atitudes. Nossas próprias atitudes e nossas sensações atuais, dizem eles, podem ser conhecidas com “quase certeza” (1977, 255; entretanto, contraste com Nisbett e Ross 1980, 200-202, que parecem simpatizar com o ceticismo de Bem sobre o acesso especial até mesmo às nossas atitudes). Gopnik admite que nós “podemos estar bem equipados para detectar certos tipos de atividade cognitiva interna de forma vaga e não especificada”, e que temos “acesso genuinamente direto e especial a certos tipos de evidência de primeira pessoa [que] podem explicar o fato de que podemos tirar algumas conclusões sobre nossos próprios estados psicológicos quando estamos perfeitamente parados e silenciosos”, embora possamos “anular essa evidência com grande facilidade” (1993a, 11-12). Ryle (1949), de forma semelhante, enfatiza a importância do comportamento externo na autoatribuição de estados mentais, enquanto reconhece a presença de “pontadas”, “emoções”, “cócegas” e até “solilóquios silenciosos”, que conhecemos em nosso próprio caso e que não parecem ser detectáveis pela observação do comportamento externo. No entanto, nenhum desses autores desenvolve uma

abordagem sobre esse autoconhecimento aparentemente mais direto. Suas teorias são, conseqüentemente, incompletas. Independentemente da importância da evidência comportamental e das teorias gerais na condução de nossas autoatribuições, diante das considerações que levam Bem, Nisbett, Wilson, Gopnik e Ryle a essas advertências, é provavelmente impossível sustentar uma visão sobre a qual há paridade completa entre atribuições de estado mental de primeira e terceira pessoa. Deve haver algum tipo de processo introspectivo ou, pelo menos, exclusivamente em primeira pessoa.

Visões de paridade eu/outro também podem ser restritas a sub-classes particulares de estados mentais: Qualquer estado mental que só pode ser conhecido por processos cognitivos idênticos aos processos pelos quais conhecemos os mesmos tipos de estados em outras pessoas, é um estado ao qual não temos acesso distintamente introspectivo. Estados para os quais a paridade é frequentemente invocada incluem traços de personalidade, motivos inconscientes, processos perceptivos primários e as bases de nossas decisões (consulte a Seção 4.2.1, abaixo para saber mais sobre isso). Aprendemos sobre esses estados em nós mesmos, talvez, quase da mesma maneira que aprendemos sobre esses estados em outras pessoas. Carruthers

(2011; veja também a Seção 4.2.2, abaixo) apresenta um caso para paridade de acesso a atitudes proposicionais como crença e desejo (em contraste com o discurso interior, representação [*imagery*] visual e similares, que ele considera introspectíveis).

2.2 Abordagens de autodetecção

Etimologicamente, o termo “introspecção” - do latim “olhar para dentro” - sugere um processo perceptivo ou quase-perceptivo. Locke escreve que temos uma faculdade de “Percepção da Operação de nossa própria Mente” que, “embora não seja o Sentido, não tendo nada a ver com Objetos externos; mesmo assim é muito parecida com ele, e poderia ser chamada, de modo suficientemente adequado, Sentido interno” (1690[1975, 105] *italico suprimido*). Kant (1782/1997) diz que temos um “sentido interno”, pelo qual aprendemos sobre aspectos mentais de nós mesmos que é paralelo, de modos importantes, ao “sentido externo”, pelo qual aprendemos sobre objetos externos.

Mas o que significa dizer que a introspecção é como a percepção? Em quais aspectos? Como Shoemaker (1994a, 1994b, 1994c) aponta, em vários aspectos, a introspecção é plausivelmente diferente da percepção. Por exemplo, a

introspecção não envolve um órgão específico como o olho ou o ouvido (embora, como observa Armstrong, 1968, nem a propriocepção corporal o faça). Tanto amigos quanto inimigos das abordagens de autodeteção tendem a concordar que a introspecção não envolve uma fenomenologia distintiva de “aparências introspectivas” (Shoemaker, 1994a, 1994b, 1994c; Lycan, 1996; Rosenthal, 2001; Siewert, 2012): a experiência visual da vermelhidão tem uma qualidade sensorial ou fenomenologia distintiva que seria difícil ou impossível de transmitir a uma pessoa cega; analogamente para a experiência olfativa de sentir o cheiro de uma banana, a uma experiência auditiva de ouvir um órgão de tubos, a experiência de tocar algo dolorosamente quente. Para ser análoga à experiência sensorial nesse respeito, a introspecção teria que gerar uma fenomenologia analogamente distintiva - alguma fenomenologia quase sensorial em adição, digamos, à fenomenologia visual do ver vermelho, que é a fenomenologia *aparição introspectiva* da fenomenologia visual do ver vermelho. Isso pareceria exigir duas camadas de aparição na percepção sensorial introspectivamente acompanhada: uma aparição visual do objeto externo e uma aparição introspectiva dessa aparição visual. (Isso não é dizer, no entanto, que a introspecção, ou pelo menos a introspecção consciente, não envolva algum tipo de “fenomenologia cognitiva” - se existe tal

coisa - do tipo que acompanha os pensamentos conscientes em geral: veja Bayne e Montague, eds., 2011.)

Os proponentes contemporâneos de modelos quase-perceptivos de introspecção concedem a existência de tais faltas de analogia (por exemplo, Lycan 1996). Podemos considerar um relato introspectivo como sendo quase-perceptivo ou, menos contenciosamente, como um relato de “autodetecção”, se atender às cinco primeiras condições descritas na Seção 1.1 - isto é, a condição de mentalidade, a condição de primeira pessoa, a condição de proximidade temporal, a condição de diretividade e a condição de detecção. Um aspecto da condição de detecção merece ênfase especial aqui: essa detecção requer a independência ontológica entre o estado mental alvo e o juízo introspectivo - os dois estados estarão causalmente conectados (assumindo que tudo ocorreu bem) mas não *constitutivamente* conectados. (Shoemaker (1994a, 1994b, 1994c) chama de modelos “perceptivos amplos” os modelos de autoconhecimento que atendem a esse aspecto da condição de detecção.) Talvez em um entendimento liberal de “detecção”, que não demande independência ontológica, abordagem de contenção ou outras abordagens da introspecção (veja a Seção 2.3.1, abaixo) possam ser qualificados como envolvendo “detecção”. Entretanto, não é assim que a “detecção” está sendo usada na presente taxonomia.

As abordagens de autodetecção do autoconhecimento parecem colocar a introspecção epistemicamente em pé de igualdade com a percepção sensorial. Para muitos filósofos, isso pareceu uma deficiência nessas abordagens. Uma longa e difundida tradição filosófica sustenta que o autoconhecimento é epistemicamente especial, que temos um “acesso privilegiado” a - talvez até um conhecimento infalível ou indubitável de - pelo menos alguma porção de nossa mentalidade, de um modo que é de tipo significativamente diferente em relação ao nosso conhecimento do mundo exterior (veja a Seção 4, abaixo). Tanto as abordagens de paridade eu/outro (Seção 2.1, acima) quanto as de autodetecção (esta seção) do autoconhecimento ~~ou~~ negam qualquer privilégio epistêmico especial ou caracterizam esse privilégio como similar ao privilégio de ser a única pessoa a ter uma visão ampliada de um objeto ou um certo tipo de acesso sensorial a esse objeto. Outras abordagens do autoconhecimento a serem discutidas mais adiante na Seção 2.3 são mais prontamente compatíveis com, e frequentemente até certo ponto conduzidas por, noções mais robustas das diferenças epistêmicas entre autoconhecimento e conhecimento de objetos ambientais.

2.2.1 Abordagens de monitoramento simples

Armstrong (1968,1981,1999) é talvez o principal defensor de uma abordagem de introspecção simples, quase perceptual, de autodeteção. Ele descreve a introspecção como um “processo de autoescaneamento no cérebro” (1968, 324), e enfatiza o que vê como a importante distinção ontológica entre o estado de consciência [*awareness*] produzido pelo procedimento de autoescaneamento e o estado mental alvo do qual a pessoa está consciente [*aware*] por meio desse escaneamento - a distinção, por exemplo, entre a dor que uma pessoa tem e a consciência introspectiva que a pessoa tem dessa dor.

Armstrong também parece sustentar que o processo introspectivo quase perceptivo procede a um nível consideravelmente baixo cognitivamente - rápido e simples, tipicamente sem muita interferência ou influência de outros processos cognitivos ou sensoriais. Ele descreve a introspecção como “completamente não inferencial”, similar à simples detecção da pressão que se tem nas costas (1968, 97) e diz que ela pode ser (e presumivelmente tipicamente é) contínua e “reflexa”, envolvendo não mais do que manter “uma breve observação sobre nossos próprios conteúdos mentais atuais, mas sem dar muita importância a isso” (1999, 115). Uma vez que

Armstrong admite que as inferências são frequentemente não-conscientes, baseadas em pistas sensoriais ou outras que a pessoa que infere não pode ela mesma discernir, sua alegação de que o processo introspectivo é não-inferencial é um compromisso substancial com a simplicidade do processo. Ele contrasta esse automonitoramento reflexivo com atos mais sofisticados de introspecção deliberada, que ele pensa também serem possíveis (1999,114). Observe que ao chamar o automonitoramento reflexivo de “introspecção”, Armstrong viola a condição de esforço da Seção 1.1, que requer que a introspecção não seja constante e automática. Lycan (1996) endossa uma visão similar, embora, diferentemente de Armstrong, Lycan caracterize a introspecção como envolvendo mecanismos *atencionais*, assim presumivelmente tratando a introspecção mais exigente de recursos cognitivos (embora talvez ainda quase constante).

Nichols e Stich (2003) empregam um modelo da mente no qual ter uma atitude proposicional tal como uma crença ou desejo é uma questão de ter uma representação armazenada em uma “caixa de crenças” ou “caixa de desejos” funcionalmente (e metaforicamente) definida (ver também os verbetes sobre crença²

² N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/belief/>>

e funcionalismo³). Na abordagem deles, a autoconsciência dessas atitudes tipicamente envolve a operação de um “Mecanismo de Monitoramento” simples que meramente pega as representações dessas caixas, acrescenta um “eu acredito que...”, “eu desejo que...”, ou qualquer outra coisa (conforme for apropriado) a essa representação e a adiciona de volta à caixa de crença. Por exemplo, se eu desejo que meu pai voe para Hong Kong no domingo, o Mecanismo de Monitoramento pode copiar a representação na minha caixa de desejos com o conteúdo “meu pai voa para Hong Kong no domingo” e produzir uma nova representação na minha caixa de crenças - isto é, criar uma nova crença - com o conteúdo “Eu desejo que meu pai voe para Hong Kong no domingo”. Nichols e Stich também propõem um mecanismo análogo, mas em certa medida mais complicado (eles deixam os detalhes não especificados) que toma perceptos como *input* e produz crenças sobre esses perceptos como *output*.

Nichols e Stich enfatizam que esse Mecanismo de Monitoramento não opera isoladamente, mas frequentemente coopera ou compete com um segundo meio de adquirir autoconhecimento, que envolve a organização ou configuração de teorias nos moldes sugeridos por Gopnik (ver Seção 2.1.2,

³ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/functionalism/>>

acima). Eles oferecem um argumento da “dupla dissociação” para essa visão. Isto é, apresentam, por um lado, casos que interpretam como casos que mostram um colapso no Mecanismo de Monitoramento, enquanto a capacidade de inferência teórica sobre a mente permanece intacta e, por outro lado, casos em que a capacidade de inferência teórica sobre a mente é prejudicada, mas o Mecanismo de Monitoramento continua a funcionar normalmente, sugerindo que a inferência teórica e o automonitoramento são processos distintos e separáveis. Nichols e Stich argumentam que pessoas autistas têm um conhecimento teórico muito pobre da mente, como sugerido pelo desempenho muito pobre delas em tarefas de “teoria da mente” (tarefas como avaliar quando alguém terá uma crença falsa), e, ainda assim, elas são bem-sucedidas em monitorar seus estados mentais como mostrado pela habilidade delas de descrever seus estados mentais em autobiografias e outras formas de autorrelato. Inversamente, Nichols e Stich argumentam que as pessoas esquizofrênicas continuam sendo excelentes teorizadores sobre estados mentais, mas monitoram muito pobremente seus próprios estados mentais - por exemplo, quando falham em reconhecer certas ações como suas próprias e têm dificuldade de relatar, ou negar a existência de pensamentos em curso.

2.2.2 Abordagens de monitoramento multiprocesso

Goldman (2006) critica a abordagem de Nichols e Stich (ver Seção 2.2.1, acima) por não descrever como o Mecanismo de Monitoramento detecta o tipo de atitude da representação (crença, desejo etc.). Se falar em “caixas de crença” e similares é um atalho para falar do papel funcional (como Nichols e Stich dizem), então o Mecanismo de Monitoramento deve, de alguma forma, detectar o papel funcional da representação detectada. Mas o papel funcional é uma questão de o que é apto a causar um estado mental particular e o que aquele estado mental é apto a causar (veja o verbete sobre funcionalismo⁴), e Goldman argumenta que um mecanismo simples não poderia discernir tais fatos de disposicionais e relacionais (embora Nichols e Stich possam ser capazes de evitar essa preocupação ao descrever a introspecção como envolvendo não apenas um, mas sim um aglomerado de mecanismos similares: 2003, 162). Goldman também argumenta que a abordagem de Nichols e Stich não deixa claro como podemos discernir a força ou intensidade de nossas crenças, desejos e outras atitudes proposicionais.

⁴ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/functionalist/>>

A abordagem positiva de Goldman começa com a ideia de que a introspecção é um processo quase perceptivo que atenciona: “A atenção parece agir como um órgão orientador na introspecção, análogo ao deslocamento do olhar ou ao farejar do nariz” (2006, 244). Os estados mentais acompanhados individualmente são então classificados em amplas categorias (de forma semelhante, na percepção visual podemos classificar objetos vistos em amplas categorias). No entanto, na visão de Goldman, esse processo só pode gerar conhecimento introspectivo dos *tipos* gerais de estados mentais (tais como crença, felicidade, sensação corporal) e algumas propriedades daqueles estados mentais (tais como grau de confiança para a crença e “uma multitude de categorias finamente delineadas” para sensação corporal). Conteúdos específicos, especialmente de atitudes como a crença, são demasiadamente diversos, sugere Goldman, para que existam categorias classificatórias pré-existentes para cada uma. Em vez disso, representamos o conteúdo específico de tais estados mentais ao reimplementar o conteúdo representacional do estado mental, isto é, simplesmente copiando o conteúdo do estado mental introspeccionado para o conteúdo da crença ou juízo introspectivo (de certa forma, como nas abordagens de Nichols e Stich). Finalmente, Goldman argumenta que alguns estados

mentais requerem “tradução” para o código mental apropriado à crença para que possam ser introspeccionados. Ele sugere que as representações visuais têm um formato ou código mental diferente das crenças e, portanto, o trabalho cognitivo será necessário para traduzir os detalhes minuciosos da experiência visual em conteúdos mentais que possam ser acreditados introspectivamente.

Hill (1991,2009) também oferece uma abordagem da introspecção de autodetecção multiprocessual. Como Goldman, Hill vê a atenção (em algum sentido amplo, não-sensorial) como central para a introspecção, embora também aceite a consciência [*awareness*] introspectiva sem atenção (1991, 117-118). Hill enfatiza as diferenças entre introspecção e percepção, ao mesmo tempo em que conserva uma abordagem amplamente de autodetecção. Hill (2009) argumenta que a introspecção é um processo que produz juízos sobre os estados alvo, em vez de consciência perceptual acerca deles, e sugere que os processos que geram esses juízos variam consideravelmente, dependendo do estado alvo, e são frequentemente complexos. Por exemplo, os juízos sobre crenças e desejos duradouros devem, diz ele, envolver procedimentos complexos para buscar armazéns de memória de longo prazo “vastos e heterogêneos”. Central para a abordagem de Hill (1991) é uma ênfase na capacidade da atenção

introspectiva de transformar - especialmente de ampliar e enriquecer, mesmo de criar - a experiência alvo. A este respeito, Hill argumenta que o ato introspectivo difere do ato observacional paradigmático que não transforma o objeto percebido (embora, é claro, tanto a observação científica quanto a ordinária - especialmente gustativa - possam afetar o que é percebido); e, assim, a abordagem de Hill contém um aspecto de “autorrealização” ou “automodelagem” no sentido da Seção 2.3.1 e da Seção 2.3.2, abaixo, e atende apenas de forma qualificada e condicional à condição de detecção em abordagens da introspecção conforme descrita na Seção 1.1, acima - a condição de que a introspecção envolva sintonização a ou detecção de um estado ou evento mental pré-existente.

Como Hill, Prinz (2004) argumenta que a introspecção deve envolver múltiplos mecanismos, dependendo tanto dos estados alvo (p. ex., atitudes versus experiências perceptivas) quanto do modo particular de acesso a esses estados. O acesso pode envolver atenção controlada ou pode ser mais um constatar passivo; pode envolver a “legendagem” ou o rotulamento verbal das experiências, ou pode envolver o tipo de acesso não-verbal que mesmo os macacos têm a seus estados mentais. Prinz (2007) distingue fortemente entre a *classificação conceitual* de nossas experiências conscientes em vários tipos que podem ser

reconhecidos e reidentificados ao longo do tempo - classificações que ele acha que devem ser necessariamente em certa medida rudimentares - e conhecimento não-conceitual de experiências conscientes em curso obtidas “ao apontar para elas com atenção. O último tipo de conhecimento, argumenta Prinz, é muito mais detalhado e finamente estruturado do que o primeiro, mas não pode ser expresso ou retido ao longo do tempo. Prinz também acompanha Hill ao enfatizar que a introspecção frequentemente intensifica ou, de outro modo, modifica a experiência do alvo. Em tais casos, argumenta Prinz, o “acesso” introspectivo é apenas acesso em um sentido atenuado.

2.3 Introspecção sem autodetecção?

Há várias maneiras de gerar juízos, ou, pelo menos, afirmações, sobre a própria vida mental atual que se tem - vamos chamá-las de autoatribuições - que são confiáveis, embora não envolvam a detecção de um estado pré-existente. Considere os quatro tipos de casos a seguir:

- A. *Autoatribuições automaticamente autorrealizadoras: Eu penso para mim mesmo: “Estou pensando”.* Ou: julgo que estou fazendo um julgamento sobre minha própria vida mental. Ou digo a mim mesmo em um discurso

interior “estou dizendo a mim mesmo em discurso interior”: “blu-bob”. Tais autoatribuições são automaticamente autorrealizadoras. Suas condições de existência são um subconjunto de suas condições de verdade.

- B. *Autoatribuições que incitam automodelagem*: Eu declaro que tenho uma imagem mental de um elefante rosa. Ao mesmo tempo em que faço essa declaração, deliberadamente causo a mim mesmo formar a imagem mental de um elefante rosa. Ou: um homem não iniciado no amor romântico declara a uma potencial namorada que ele é o tipo de pessoa que envia flores para suas amadas. Ao mesmo tempo em que ele diz isso, ele decide com sucesso ser o tipo de pessoa que envia flores para suas amadas. A autoatribuição ou precipita uma mudança ou fortalece o que já existe, de forma a tornar a autoatribuição precisa. Nestes casos, diferentemente dos casos descritos em (A), alguma mudança ou automanutenção é necessária para tornar a autoatribuição verdadeira, além do evento de autoatribuição por si só.
- C. *Autoatribuição precisa por meio da autoexpressão*: Eu aprendo a dizer “Estou com dor!” em vez de “ai!” como

uma resposta automática e irrefletida a estímulos dolorosos. Ou: uso a sentença autoatributiva “Acredito que Russell mudou de ideia sobre pacifismo” simplesmente como uma forma cautelosa de expressar a crença de que Russell mudou de ideia sobre pacifismo, sendo esta expressão o produto do refletir sobre Russell, em vez de um produto da reflexão sobre minha própria mente. Presume-se aqui que autoexpressões deste tipo fluem naturalmente dos estados expressos mais ou menos da mesma forma que expressões faciais e expressões verbais não autoatributivas fluem naturalmente desses mesmos estados – isto é, sem serem precedidas por qualquer tentativa de detectar o estado autoatribuído.

D. *Autoatribuições derivadas de julgamentos sobre o mundo exterior*: Do fato não autoatributivo de que Stanford fica ao sul de Berkeley, eu derivo a conclusão autoatributiva de que eu acredito que Stanford fica ao sul de Berkeley. Ou: do fato não autoatributivo de que seria bom ir para casa agora, eu derivo o julgamento autoatributivo de que quero ir para casa agora. Essas derivações podem ser inferências, mas, se assim for, tais inferências não requerem nenhuma premissa específica sobre estados mentais em andamento. Talvez se adote um princípio

geral de inferência como “de *P*, é permissível derivar que eu *acredito que P*”, ou “normalmente, se algo é bom, eu o quero.”

Todas as seguintes abordagens de autoconhecimento tiram proveito de um ou mais desses fatos sobre a autoatribuição. Porque todas essas formas de obter autoconhecimento violam a condição de detecção na introspecção (condição 5 na Seção 1.1, acima), e porque os filósofos estão divididos sobre se os métodos de obter autoconhecimento que violam esta condição contam, estritamente falando, como métodos introspectivos, os filósofos estão divididos sobre se as abordagens de autoconhecimento do tipo descrito nesta seção deveriam ser considerados como abordagens de introspecção.

2.3.1 Autorrealização e contenção

Uma ênfase no conhecimento infalível através de autoatribuições autorrealizadoras remonta pelo menos a Agostinho (c. 420 C.E./1998) e é mais conhecidamente implementada por Descartes em seu *Discurso do Método* (1637/1985) e *Meditações* (1641/1984), nos quais ele assume o pensamento autorrealizador de que ele está pensando como indubitavelmente verdadeiro, imune até mesmo ao ceticismo

mais radical, e um fundamento seguro sobre o qual se pode construir conhecimentos adicionais (pelo menos sob certas condições; veja Paul 2018).

As abordagens contemporâneas de autorrealização tendem a explorar a ideia de *contenção*. Em um ensaio de 1988, Burge escreve:

Quando alguém sabe que está pensando em p, não está tomando o seu pensamento (ou “pensando”) neste p meramente como um objeto. Pensa-se no p no próprio “evento” de pensar, com conhecimento de causa, que está pensando isto. Isto é pensado e é pensamento no mesmo ato mental (654).

Esse é o caso, Burge argumenta, porque “por seu caráter reflexivo e autorreferencial, o conteúdo do julgamento de segunda ordem [autoatributivo] está preso (autorreferencialmente) ao conteúdo de primeira ordem que ele tanto contém quanto toma como seu assunto” (1988, 659-660; cf. Heil 1988; Gertler 2000, 2001; Heil e Gertler descrevem tais pensamentos como introspectivos enquanto Burge parece não pensar no autoconhecimento assim construído como introspectivo: 1998, 244; ver também 1988, 652). Ao julgar que estou pensando em uma banana, com isso, penso

necessariamente em uma banana: o julgamento autoatributivo contém, como uma parte, o próprio pensamento autoatribuído, e, portanto, não pode ser falso. Em um ensaio de 1996, Burge estende suas observações para incluir não apenas “pensamentos” autoatributivos como alvos, mas também (certos tipos de) “julgamentos (por exemplo, “Eu julgo, com isto, que existem entidades físicas” e outros juízos “com reflexividade semelhante a isto”, 92).

Shoemaker (1994a, 1994b, 1994c) emprega a ideia de contenção de forma muito diferente e sobre uma gama muito mais ampla de alvos introspectivos, incluindo estados conscientes como dores, e atitudes proposicionais, como crenças. Shoemaker especula que a relação relevante de contenção não se mantém entre os *conteúdos* ou *conceitos empregados* no estado alvo e no estado autoatributivo, mas sim entre suas realizações neurais no cérebro. Para desenvolver este ponto, Shoemaker distingue entre a "realização nuclear" de um estado mental e sua "realização total". Pode-se pensar em processos mentais como acontecendo em regiões consideravelmente estreitas do cérebro (sua realização nuclear), e ainda assim, sugere Shoemaker, não é como se pudéssemos simplesmente cortar fora essas regiões de todas as outras e ainda ter o estado mental em questão. Para ser o estado mental que ele é, o processo tem de estar situado em uma rede

causal maior que envolve mais do cérebro (a realização total). As relações de contenção ou sobreposição entre a realização nuclear e a realização total entre o estado alvo e o julgamento autoatributivo podem, então, subscrever a precisão introspectiva. Por exemplo, a realização total do estado cerebral do estado de dor pode ser simplesmente um subconjunto da realização total do estado cerebral do estado de acreditar que se está com dor. A precisão introspectiva poderia então ser explicada pelo fato de que o juízo introspectivo não é um estado que existe de forma independente.

Mais recentemente, os filósofos aplicaram modelos de contenção de conteúdo similares ao de Burge (em oposição aos modelos de contenção de realização similares ao de Shoemaker) ao autoconhecimento de estados conscientes, ou “fenomenologia”, em particular - por exemplo, Gertler (2001), Papineau (2002), Chalmers (2003), Horgan e Kriegel (2007), e Balog (2012). Husserl (1913/1982) oferece uma abordagem inicial de contenção fenomenal, argumentando que podemos a qualquer momento colocar nossa “*cogitatio*” - nossas experiências conscientes - conscientemente diante de nós por meio de um tipo de olhar mental, com a autopercepção que surge contendo como uma parte a experiência consciente para a qual ela é dirigida e incapaz de existir sem. Papineau oferece uma abordagem

“citacional” na qual, na introspecção, nós nos autoatribuímos “a experiência: ___”, na qual o espaço em branco é completado pela experiência ela mesma. Chalmers escreve que “crenças fenomenais diretas” sobre nossas experiências são “parcialmente *constituídas* por uma qualidade fenomenal subjacente”, na medida em que as duas serão firmemente acopladas em “uma ampla variedade de casos possíveis conceitualmente próximos” (2003, 235).

Uma possível dificuldade com tais abordagens é que, enquanto pareça plausível supor que um pensamento ou julgamento introspectivo possa conter outro pensamento ou julgamento como uma parte, é menos claro como um juízo ou crença autoatributivos poderiam conter um pedaço da experiência consciente como uma parte. Crenças, e outros estados mentais tais quais crenças, como julgamento, pode-se pensar, contêm *conceitos*, não experiências conscientes, como seus constituintes (Fodor 1998); ou, alternativamente, pode-se pensar que crenças são padrões funcionais ou disposicionais de resposta ao input (Dennett 1987; Schwitzgebel 2002), mais uma vez tornando incerto como um pedaço da fenomenologia poderia ser parte da crença. Talvez com esta preocupação em mente, os defensores das abordagens de contenção frequentemente apelam para “conceitos fenomenais” que são, como os julgamentos

introspectivos para os quais eles contribuem, parcialmente constituídos pelas experiências conscientes que são os conteúdos destes conceitos. Pensa-se frequentemente que tais conceitos são obtidos por meio da atenção demonstrativa às nossas experiências conscientes na medida em que elas estão em curso.

Pareceria, pelo menos, que crenças, conceitos ou julgamentos contendo pedaços de fenomenologia teriam que expirar uma vez que a fenomenologia tenha passado e, assim, que os julgamentos introspectivos não poderiam ser usados em inferências posteriores sem recriar o estado em questão. Chalmers (2003) concede a localidade temporal de tais juízos introspectivos contendo fenomenologia e, conseqüentemente, seu uso limitado na fala e na realização de generalizações. Papineau (2002), em contraste, adota uma teoria na qual a recriação imaginativa da fenomenologia no pensar sobre a experiência passada é lugar-comum.

2.3.2 Automodelagem

Embora aparentemente possamos, pelo menos às vezes, chegar a verdadeiras autoatribuições por meio dos procedimentos de automodelagem e autoexpressão (B e C) descritos no início da Seção 2.3, e embora tais procedimentos possam atender às três primeiras condições para uma abordagem

de introspecção, conforme descrito na Seção 1.1 - isto é, podem (dependendo de como são descritos e desenvolvidos) ser procedimentos que podem produzir apenas conhecimento ou julgamentos (ou pelo menos autoatribuições) sobre processos mentais, sejam eles atuais ou ocorridos muito recentemente - poucos filósofos descreveriam tais procedimentos como “introspectivos”. Ainda assim, eles merecem aqui um tratamento breve, parcialmente pela mesma razão que as abordagens de paridade eu/outro mereceram tratamento na Seção 2.1, acima - isto é, como abordagens céticas que sugerem que o escopo da introspecção pode ser consideravelmente mais restrito do que geralmente se pensa - e em parte como pano de fundo para as abordagens de “transparência” a serem discutidas na Seção 2.3.4, abaixo, com as quais eles são frequentemente casados.

É difícil encontrar abordagens de autoconhecimento que enfatizem a técnica de automodelagem em sua forma mais pura, prospectiva e causal - talvez porque seja claro que o autoconhecimento tem de envolver consideravelmente mais do que isso (Gertler 2011). No entanto, McGeer (1996, 2008; McGeer e Pettit 2002) coloca considerável ênfase na automodelagem, escrevendo que “nós aprendemos a usar nossas autoatribuições intencionais para instilar ou reforçar tendências e

inclinações que se adequam a essas atribuições, mesmo que tais tendências e inclinações possam, no melhor dos casos, ter sido apenas nascentes no momento em que nós inicialmente fizemos os juízos” (1996, 510). Se eu me descrevo como valente em batalha, ou como um vegetariano comprometido- especialmente se o faço publicamente- crio compromissos e expectativas para mim mesmo que ajudam a tornar essas autoatribuições verdadeiras. McGeer compara o autoconhecimento com o conhecimento que um motorista tem, em oposição ao de um passageiro, sobre para onde o carro está indo: o motorista, diferentemente do passageiro, pode fazer com que seja o caso que o carro vá para onde ele diz que está indo (505).

Há também tensões em Dennett (embora Dennett possa não ter uma visão inteiramente consistente sobre estes assuntos; veja Schwitzgebel 2007) que sugerem uma visão de autorrealização ou de automodelagem. Em alguns lugares, Dennett compara autorrelatos “introspectivos” sobre a consciência a obras de ficção, imunes à refutação da mesma forma que reivindicações ficcionais o são - não se poderia errar mais sobre sua própria consciência, diz Dennett, do que Doyle poderia errar sobre a cor da poltrona de [Sherlock] Holmes (por exemplo, 1991, 81, 94). Tais observações são consistentes ou

com uma visão antirrealista da ficção (não há fatos sobre a poltrona ou sobre a consciência; ver 366-367) ou com uma visão realista de autorrealização ou automodelagem (Doyle *cria* fatos sobre Holmes enquanto pensa ou escreve sobre ele; nós criamos fatos sobre como é sermos nós mesmos ao pensar ou fazer reivindicações sobre nossa consciência, como talvez em 81 e 94). Mais moderadamente, ao discutir atitudes, Dennett enfatiza como o ato de formular uma atitude em linguagem - por exemplo, ao pedir um item do menu - pode envolver autoatribuir um grau de especificação nas atitudes que se tem que não estava presente antes, com isso, comprometendo a pessoa a, e criando parcialmente ou totalmente, a atitude específica autoatribuída (1987, 20).

Abordagens *comissivas* do autoconhecimento também envolvem automodelagem, mas não uma forma de automoldagem na qual o julgamento introspectivo traz à existência um estado alvo ontologicamente distinto, mas sim um tipo de automodelagem envolvendo um componente de autorrealização ou contenção similar ao discutido na Seção 2.3.1, acima. Moran (2001), por exemplo, argumenta que normalmente quando somos levados a pensar no que acreditamos, desejamos ou pretendemos (e ele limita sua abordagem primariamente a

estes três estados mentais), refletimos sobre os fenômenos (externos) em questão e decidimos o que acreditar, desejar ou fazer. Em vez de tentar detectar um estado pré-existente, abrimos ou reabrimos o assunto e chegamos a uma resolução. Dado que normalmente nós, de fato, acreditamos, desejamos e pretendemos o que resolvemos acreditar, desejar e fazer, podemos, portanto, autoatribuir com precisão essas atitudes. Coliva (2016) argumenta que a autoatribuição “eu acredito que P” é como uma declaração performativa na medida em que constitui um comentário à crença de que P [seja o caso] (ver também Wright 1989; Falvey 2000; Heal 2002; Boyle 2009, no prelo.)

2.3.3 Expressivismo

Wittgenstein escreve:

[C]omo um ser humano aprende o significado dos nomes das sensações? - da palavra “dor”, por exemplo. Aqui está uma possibilidade: as palavras estão ligadas às expressões primitivas, naturais da sensação e usadas em seu lugar. Uma criança se machucou e chora: então os adultos conversam com ela e a ensinam exclamações e, mais tarde, frases. Eles ensinam à criança um novo comportamento-dor.

“Então você está dizendo que a palavra ‘dor’ realmente significa chorar?” - Pelo contrário: a expressão verbal de dor substitui chorar e não o descreve. (1953/1968, sec. 244)

E:

Não se pode dizer de mim (exceto talvez como uma brincadeira) que sei que estou com dor. O que isso deveria significar, exceto talvez que estou com dor? (1953/1968, sec. 246).

Do ponto de vista de Wittgenstein, é tanto verdade que estou com dor quanto que digo de mim mesmo que estou com dor, mas a enunciação não emerge de forma alguma de um processo de *detecção* da dor.

Uma visão expressivista simples - às vezes atribuída a Wittgenstein com base nestas e em outras passagens relacionadas - nega que as enunciações expressivas (por exemplo, “isso dói!”) genuinamente atribuem estados mentais aos indivíduos que as pronunciam. Tal visão enfrenta sérias dificuldades para acomodar as semânticas evidentes das enunciações autoatributivas, incluindo seu uso na inferência e as simetrias aparentes entre os usos do tempo presente e do tempo passado e entre os usos da primeira e da terceira pessoas (Wright 1998; Bar-

On 2004). Consequentemente, Bar-On defende, ao invés disso, o que ela chama de visão neoexpressivista, segundo a qual enunciações expressivas podem compartilhar estrutura lógica e semântica com enunciações não-expressivas, apesar das diferenças epistêmicas entre elas.

Expressivistas nem sempre foram claros sobre o alcance exato dos estados mentais alvo desta forma, mas parece plausível que, pelo menos a princípio, algumas autoatribuições verdadeiras (ou aptas) poderiam surgir deste modo, sem nenhuma intervenção de autodeteção introspectiva. A questão então seria se é assim que nós *geralmente* chegamos a autoatribuições verdadeiras, para alguma classe particular de estados mentais, ou se algum processo mais arquetipicamente introspectivo também está disponível. (Para um tratamento mais detalhado do expressivismo, consulte a seção sobre o modelo expressivista do autoconhecimento no verbete autoconhecimento⁵).

2.3.4 Transparência

Evans escreve:

⁵ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/self-knowledge/>>

Ao fazer uma autoatribuição de crença, os olhos são, por assim dizer, ou às vezes literalmente, dirigidos para fora – sobre o mundo. Se alguém me pergunta “Você acha que haverá uma terceira guerra mundial?”, ao responder-lhe, eu devo atentar precisamente para os mesmos fenômenos externos aos quais eu atentaria se estivesse respondendo à pergunta “Haverá uma terceira guerra mundial?” Coloco-me na posição de responder à questão se eu acredito que p (1982, 225).

Abordagens de *transparência* ao autoconhecimento, como a de Evans, enfatizam casos em que parece que se chega a uma autoatribuição precisa, não por meio de assistir ou pensar sobre os próprios estados mentais, mas sim por meio de assistir ou pensar sobre estados externos do mundo aos quais os estados mentais alvos se referem. Note que esta afirmação tem tanto um aspecto negativo quanto positivo: *não* aprendemos sobre nossas mentes, por assim dizer, olhando para dentro; e nós *aprendemos* sobre nossas mentes refletindo sobre os aspectos do mundo aos quais nossos estados mentais se referem. As teses positivas e negativas são separáveis: um pluralista poderia aceitar a tese positiva sem a negativa; um defensor da teoria de paridade eu/outro ou de uma abordagem expressivista do autoconhecimento (com respeito a uma certa classe de estados

alvo) poderia aceitar a tese negativa sem a positiva. (Note bem: na literatura filosófica sobre autoconhecimento, “transparência” às vezes também é usada para referir-se a algo como autointimação no sentido da seção 4.1.1, abaixo, por exemplo em Wright 1998; Bilgrami 2006. Esse é um uso completamente diferente, que não deve ser confundido com o uso presente). Porque as abordagens da transparência enfatizam o foco externo do nosso pensamento ao chegar a autoatribuições, chamar tais abordagens de abordagens de “introspecção” tensiona contra a etimologia do termo. Entretanto, alguns proeminentes defensores das abordagens de transparência, tais como Dretske (1995) e Tye (2000), as apresentam explicitamente como abordagens de introspecção.

A variedade de estados alvos aos quais a transparência se aplica é tema de certa disputa. Entre os filósofos que aceitam algo como semelhante a transparência, a crença é geralmente considerada como transparente (Gordon 1995, 2007; Gallois 1996; Moran 2001; Fernández 2003; Byrne 2018). Os estados ou experiências perceptivas também são frequentemente considerados como transparentes no sentido relevante. O exemplo de Harman é o mais citado:

Quando Eloise vê uma árvore diante dela, as cores que ela experimenta são todas experienciadas como características da árvore e de seu entorno. Nenhuma delas é experienciada como características intrínsecas de sua experiência. Ela também não experimenta quaisquer características de nada como características intrínsecas de suas experiências. E isso também é verdadeiro para você também. Não há nada de especial na experiência visual de Eloise. Quando você vê uma árvore, você não experimenta nenhuma característica como característica intrínseca da sua experiência. Olhe para uma árvore e tente voltar sua atenção para características intrínsecas da sua experiência visual. Prevejo que você descobrirá que as únicas características lá para os quais você pode voltar sua atenção serão as características da árvore apresentada. (Harman 1990, 667).

A ênfase de Harman aqui está na tese negativa, que remonta, pelo menos, a Moore (1903; embora Moore não a endosse inequivocamente). A visão de que é impossível observar diretamente à experiência perceptiva tem sido especialmente enfatizada por Tye (1995, 2000, 2002; veja também Evans 1982; Van Gulick 1993; Shoemaker 1994a; Dretske 1995; Martin 2002; Stoljar 2004), e conflita diretamente com abordagens segundo as

quais aprendemos sobre nossa experiência sensorial, principalmente ao direcionar atenção introspectiva a ela (por exemplo, Goldman 2006; Petitmengin 2006; Hill 2009; Siewert 2012; e remonta, pelo menos, à Wundt 1888 e Titchener 1908 [1973]).

Gordon (2007) argumenta (contra Nichols e Stich 2003 e Goldman 2006) que as *rotinas ascendentes* similares às de Evans (ascendendo de “p” para “eu acredito que p”) podem conduzir autoatribuição precisa de todas as atitudes, não apenas da crença. Ele defende sua proposta por meio do casamento da tese da transparência com algo como uma teoria expressiva da autoatribuição: para responder a uma pergunta sobre o que eu quero - por exemplo, qual sabor de sorvete eu quero? - eu não penso nos meus desejos, mas antes nos diferentes sabores disponíveis, e então eu *expresso* a atitude autoatributivamente. De forma semelhante para esperanças, medos, desejos, intenções, arrependimentos etc. Gordon aponta que desde muito cedo, antes que provavelmente tenham qualquer intenção autoatributiva, as crianças aprendem a expressar suas atitudes autoatributivamente, por exemplo, com frases simples como “[Eu] quero banana!” (veja também Bar-On 2004).

As abordagens comissivas de autoconhecimento (ver seção 2.3.2, acima) geralmente também afirmam transparência: refletir

sobre o mundo gera compromisso com uma crença, desejo ou intenção, que alguém, com isso, também conhece ou se autoatribui (Falvey 2000; Moran 2001; Coliva 2016; Boyle, no prelo).

A tese de transparência é de fato consistente, não apenas com abordagens expressivistas e comissivas, mas com qualquer um dos quatro procedimentos de autoatribuição não baseados em detecção descritos no início desta seção (e, de fato, Aydede e Güzeldere tentam reconciliar aspectos da visão de transparência com uma abordagem à introspecção amplamente similar à de detecção). Essa múltipla compatibilidade decorre do fato de que, por si só, a tese de transparência, por si só, não vai muito longe em direção a uma visão positiva dos mecanismos de autoconhecimento.

Byrne (2018) e Dretske (1995) reúnem transparência e algo como um modelo derivacional de autoconhecimento - um modelo no qual eu derivo a conclusão de que acredito que P diretamente do próprio P , ou a conclusão de que estou representando x como F a partir do fato de que x é F - um fato que, para servir como premissa na derivação, deve, naturalmente, ser representado (ou acreditado) por mim. Byrne argumenta que, assim como se pode tolerar a seguinte regra epistêmica:

CAMPAINHA: Se a campainha tocar, acredite que há alguém à porta.

Então alguém também poderia tolerar a regra:

BEL: Se P , acredite que você acredita que P .

Para determinar se você acredita em P , primeiro determine se P é o caso, então siga a regra BEL. Byrne (2018) oferece abordagens similares do autoconhecimento de intenção, pensamento, visão e desejo.

Dretske faz uma analogia da introspecção com casos ordinários de “percepção deslocada” - casos em que se percebe que algo é o caso mediante a percepção direta de alguma outra coisa (por exemplo., escutar que o carteiro chegou ao ouvir o cão ladrar; ver que você pesa 50 quilos ao olhar o mostrador na balança do banheiro): percebe-se que se representa x como F por meio da percepção da F -dade de x . Dretske observa, entretanto, dois pontos de ausência de analogias entre os casos. No caso de escutar que o carteiro chegou ao ouvir o cão ladrar, a conclusão (de que o carteiro chegou) só é estabelecida se a premissa sobre o latido do cão for verdadeira e, além disso, depende de uma crença de conexão revogável de que o latido do cão é um indicador confiável da chegada do correio. No caso introspectivo,

entretanto, a inferência, se é uma inferência, não requer a verdade da premissa de que x ser F . Mesmo que x não seja F , a conclusão de que estou representando x como F é apoiada. Tampouco parece haver qualquer tipo de crença de conexão revogável.

Tye também enfatiza a transparência em sua abordagem da introspecção, embora ele limite suas observações à introspecção da experiência consciente, ou “caráter fenomenal”. Em seu livro de 2000, Tye desenvolve uma visão como a de Dretske, fazendo uma analogia da introspecção com a percepção deslocada, embora Tye, ao contrário de Dretske, negue explicitamente que a inferência esteja envolvida, propondo, em vez disso, um mecanismo similar ao tipo de mecanismo previsto por abordagens de monitoramento simples como aquelas de Nichols e Stich (2003; veja a seção 2.2.1, acima), um processo confiável que, no caso da autoconsciência [*self-awareness*] perceptual toma a consciência [*awareness*] de coisas externas como seu input e produz como seu output a consciência [*awareness*] de caráter fenomenal. (A diferença-chave entre a abordagem de Tye de 2000, de um lado, e a abordagem de Nichols e Stich, do outro, que justifica a qualificação da visão de Tye aqui, mais do que na seção sobre modelos de autodetecção, é esta: Tye rejeita a ideia de que o processo é de detecção interna, enquanto Nichols e Stich enfatiza essa ideia. Para julgar a disputa entre essas duas

posições, e para determinar se ela poderia, de fato, ser meramente nominal, seria útil ter uma ideia mais clara do que tem sido apresentado até agora sobre o que significa dizer que um sistema subpessoal detecta, ou “monitora” ou “escaneia”, os estados ou o conteúdo de outro.) Contudo, em seu livro de 2009, Tye rejeita o modelo de percepção deslocada em favor de uma versão da visão de transparência que *identifica* o caráter fenomenal com qualidades externas no mundo, de modo que perceber características do mundo seja apenas perceber o caráter fenomenal - uma visão que ele reconhece ser então encarregada da difícil tarefa de explicar como o caráter fenomenal é uma propriedade (ou “qualidade”) de objetos externos em vez de, como é geralmente assumido, uma propriedade apenas de experiências desses objetos.

Diversos autores desafiaram a ideia de que a experiência sensorial necessariamente escapa à atenção - isso é, eles negaram a reivindicação central das teorias de transparência sobre a experiência sensorial. Block (1996), Kind (2003) e Smith (2008) argumentaram que fosfenos - aquelas pequenas luzes que você vê quando pressiona seus olhos - e o embaçamento visual são aspectos das experiências sensoriais que podem ser atentados diretamente. Siewert (2004) argumentou que o que é intuitivamente atraente na visão de transparência é

principalmente a observação de que, ao refletir sobre a experiência sensorial, *não se retira* a atenção dos objetos sentidos; mas, ele argumenta, isso é compatível com dedicar, também, um certo tipo de atenção à própria experiência sensorial. Nas primeiras discussões sobre atenção, a atenção perceptiva era às vezes diferenciada da “atenção intelectual” (James 1890 [1981]; Baldwin 1901-1905; veja também Peacocke 1998; Mole 2011), isto é, do tipo de atenção que podemos dedicar à quebra-cabeças de palavras puramente imaginadas ou a questões filosóficas. Se formas não sensoriais de atenção forem possíveis, então a tese de transparência para a experiência sensorial requererá uma reformulação: é apenas a atenção sensorial à experiência sensorial que é impossível? Ou é qualquer tipo de atenção? Simplesmente dizer que não atentamos sensorialmente a nossos estados mentais é fazer apenas uma reivindicação modesta, similar à reivindicação de que vemos objetos em vez de vermos nossas experiências visuais de objetos; mas dizer que não podemos atentar a nossos estados mentais, mesmo intelectualmente, parece extremo. À luz disso, permanece incerto como moldar a intuição da transparência para melhor trazer à tona a ideia central que deve ser transmitida pelo lema de que a introspecção da experiência sensorial não é uma questão de atentar à própria mente.

2.4 Pluralismo introspectivo

Filósofos discutindo o autoconhecimento frequentemente escrevem como se abordagens que destacam um desses métodos, que não envolvem autodeteção, de geração de autoatribuições conflitassem com abordagens que destacam outros desses métodos que não envolvem autodeteção, e como se abordagens desse tipo geral conflitassem com abordagens de autodeteção (seção 2.2, acima). Embora certamente existam conflitos entre diferentes abordagens destinadas a servir como abordagens *exaustivas* do autoconhecimento, é implausível que qualquer uma ou mesmo algumas destas abordagens do autoconhecimento seja *exaustiva*. Plausivelmente, todas as abordagens que não envolvem autodeteção descritas acima podem levar, pelo menos ocasionalmente, a autoatribuições precisas. Entusiastas de outros dos modelos, ou de um modelo de autodeteção, não precisam negar isso. Também parece difícil negar que, pelo menos *às vezes*, chegamos a conclusões sobre nossas vidas mentais com base no tipo de inferência teórica ou autointerpretação enfatizada por defensores da abordagem de paridade eu/outro (seção 2.1, acima). Finalmente, mesmo os filósofos preocupados com visões fortes ou excessivamente simplistas de autoescaneamento podem querer garantir que a mente possa fazer *algum* tipo de

rastreamento de seus próprios estados presentes ou ocorridos recentemente - por exemplo, quando rastreamos um fluxo de pensamentos ocorridos recentemente que, presumivelmente, não podem (porque é passado) ser autoatribuído por métodos de autorrealização, automodelagem, autoexpressão ou transparência.

Schwitzgebel (2012) eleva este pluralismo a uma espécie de abordagem negativa da introspecção. Os julgamentos introspectivos, diz ele, surgem de uma confluência mutável de muitos processos, recrutados oportunisticamente, nenhum dos quais pode ser chamado de introspecção propriamente dita. Assim como não existe uma capacidade singular, unificada de percepção de pôsteres que se empregue quando se tenta perceber as informações em um pôster em uma conferência psicológica ou feira de ciência, não existe, na visão de Schwitzgebel, nenhuma capacidade singular, unificada de introspecção ou um processo central subjacente, nem mesmo alguns poucos processos ou mecanismos dedicados. Em vez disso, o introspectador, como o observador de pôsteres, vem a empregar uma gama diversificada de recursos cognitivos, como convém à ocasião. Embora, diz ele, não valha a pena um processo ser chamado de "introspectivo", a menos que o introspectador pretendesse alcançar um juízo sobre sua experiência consciente atual ou ocorrida muito recentemente,

usando, pelo menos, alguns recursos específicos para o caso de primeira pessoa e que exibissem alguma sensibilidade relativamente direta ao estado alvo, esta limitação não implica a existência de nenhum processo introspectivo dedicado. Os defensores de versões menos extremas do pluralismo que são compatíveis com a existência de vários processos introspectivos específicos incluem Prinz (2004), Hill (2009), Coliva (2016) e Samoilova (2016).

3. O Papel da Introspecção na Psicologia Científica

3.1 A ascensão da psicologia introspectiva como ciência

Os filósofos há muito tempo fazem reivindicações introspectivas sobre a mente humana - ou, para falar com mais cautela, eles fizeram alegações aparentemente, pelo menos em parte, fundamentadas introspectivamente. Aristóteles (3º S. AEC/1961) afirma que o pensamento não ocorre sem imagens. Mengzi (3º S. AEC/2008) argumenta que nossos corações estão satisfeitos com a bondade moral e revoltados com o mal, mesmo que o prazer e a repulsa não sejam evidentes em nosso comportamento exterior. Berkeley não encontra em si “ideias abstratas” como a de um triângulo que é, nos termos de Locke, “nem oblíquo, nem retângulo, nem equilátero, isósceles, nem escaleno, mas tudo e nada disso ao mesmo tempo” (Berkeley 1710/1965, 12; Locke 1689/1975, 596). James Mill (1829[1878]) tenta um catálogo das variedades da experiência sensorial.

Embora vários dos primeiros filósofos modernos tivessem como objetivo iniciar o estudo científico da mente, não foi até meados do século XIX - com o aparecimento de *métodos quantitativos introspectivos*, especialmente no que diz respeito à consciência sensorial - que o estudo da mente tomou forma como

uma ciência progressiva, matemática, baseada em laboratório. Os primeiros psicólogos quantitativos, como Helmholtz (1856/1962), Fechner (1860[1964]) e Wundt (1896[1902]) buscaram respostas quantitativas para perguntas como: em quanto devem diferir dois estímulos físicos para que suas experiências sejam notadamente diferentes? Quão fraco um estímulo ainda pode ser percebido conscientemente? Qual é a relação matemática entre intensidade do estímulo e a intensidade da sensação resultante? (A lei Weber-Fechner sustenta que a relação é logarítmica.) Ao longo de quais dimensões, exatamente, a experiência sensorial pode variar? (A “cor sólida” [veja o link para o sólido de Munsell na nota de rodapé⁶], por exemplo, caracteriza a experiência de cor pelo apelo a apenas três dimensões de variação: matiz, saturação e luminosidade ou brilho). Embora desde muito cedo os psicólogos também empregassem métodos não-introspectivos (por exemplo, desempenho em testes de memória, tempos de reação), a maioria das caracterizações iniciais do campo colocava a introspecção no centro. James, por exemplo, escreveu que “a observação introspectiva é o que temos que confiar em primeiro lugar e sempre” (1890[1981,185]).

⁶ N. do T. Ver em <
http://facweb.cs.depaul.edu/sgrais/munsell_color_system.htm>

Em contraste com a tradição filosófica dominante que, desde Descartes, enfatizou o privilégio especial ou pelo menos a alta precisão dos juízos introspectivos sobre a consciência (veja a Seção 4.1, abaixo), muitos primeiros psicólogos introspectivos sustentavam que a introspecção da experiência consciente atual ou do passado recente é difícil e propensa a erros se o observador introspectivo for insuficientemente treinado. Wundt, por exemplo, supostamente não deu crédito aos relatos introspectivos de pessoas com menos de 50.000 testes de prática na observação de sua experiência consciente (Boring 1953). Titchener, um importante psicólogo introspectivo americano, escreveu um manual de treinamento introspectivo de 1600 páginas para estudantes, argumentando que a observação introspectiva é pelo menos tão difícil quanto a observação nas ciências físicas (Titchener 1901-1905; ver também Wundt 1874[1908]; Müller 1904; para discussões contemporâneas sobre treinamento introspectivo ver Varela 1996; Nahmias 2002; Schwitzgebel 2011b). Essa diferença de otimismo sobre a introspecção não treinada pode refletir em parte diferenças nos tipos de juízos em primeiro plano nas duas disciplinas. Os filósofos enfatizando seu privilégio tendem a se concentrar em juízos grosseiros e (aparentemente) simples, como “estou tendo uma experiência visual de vermelhidão” ou “acredito que está chovendo”. Os

projetos de interesse dos psicólogos introspectivos frequentemente exigiam juízos muito mais refinados - tais como determinar com precisão matemática se uma sensação visual tem o dobro da “intensidade” de outra ou determinar em quais dimensões a experiência emocional pode variar.

3.2 O ceticismo inicial sobre a observação introspectiva

As discussões teóricas dos primeiros psicólogos introspectivos sobre a natureza da introspecção eram frequentemente enquadradas em reação ao ceticismo sobre a viabilidade científica da introspecção, especialmente a preocupação de que o ato introspectivo interfira ou destrua o estado ou processo mental que é seu alvo⁷. A formulação mais influente dessa preocupação foi a de Comte:

⁷ Uma preocupação relacionada na discussão contemporânea é a ilusão da luz da geladeira. A ilusão da luz da geladeira tem esse nome devido ao erro que uma criança pode cometer ao pensar que a luz da geladeira está sempre acesa uma vez que está acesa sempre que ela olha para ver se está. Somos vítimas dessa ilusão, como por vezes é sugerido, se pensarmos que temos a experiência visual constante da periferia do campo visual ou a experiência tátil constante de nossos pés em nossos sapatos. Nesta visão, não temos tais experiências periféricas constantes, como sugerido, mas sempre que pensamos ou introspectamos sobre tais assuntos, criamos tais experiências, resultando na ilusão de que elas estão sempre presentes. Veja Jaynes, 1976; Dennett, 1991; Tomás, 1999; Noé, 2004; Bloco, 2007; Schwitzgebel, 2007a. Entre aqueles que sustentam que temos experiências periféricas constantes estão James, 1890/1981, e Searle, 1992.

Mas quanto a observar da mesma maneira fenômenos intelectuais no momento de sua presença real, isso é uma impossibilidade manifesta. O pensador não pode dividir-se em dois, dos quais um raciocina enquanto o outro o observa seu raciocinar. Sendo o órgão observado e o órgão observando, neste caso, idênticos, como poderia ocorrer a observação? Este pretenso método psicológico é então radicalmente nulo e sem efeito (1830, usando a tradução de James 1890 [1981,188])

Os psicólogos introspectivos tendiam a reagir a essa preocupação de três maneiras. A abordagem mais concessiva - recomendada, por exemplo, por James (1890 [1981]; ver também Mill, 1865 [1961]; Lyons, 1986) - era conceder o ponto de Comte a favor da introspecção *concorrente*, isto é, introspecção simultânea com o estado ou processo alvo e enfatizar em contraste a *retrospecção imediata*, isto é, refletir ou atentar ao processo alvo (geralmente uma experiência consciente) muito brevemente *após* ele ocorrer. Como a observação científica ocorre somente após a conclusão do processo alvo, ela não interfere nesse processo; mas é claro que o atraso entre o processo e a observação deve ser o mais breve possível para garantir que o processo seja lembrado precisamente.

Brentano (1874 [1973]) respondeu à preocupação de Comte distinguindo entre “observação interna” [*innere Beobachtung*] e “percepção interna” [*innere Wahrnehmung*]. Observação, como Brentano a caracteriza, envolve dedicar toda atenção a um fenômeno, com o objetivo de apreendê-lo precisamente. Essa dedicação de atenção interfere necessariamente no processo a ser observado se o processo for mental; portanto, diz ele, a observação interior é problemática como método psicológico científico. *Percepção* interior, em contraste, de acordo com Brentano, não envolve atenção às nossas vidas mentais e, portanto, não as perturba indesejavelmente. Enquanto nossa “atenção está voltada para um objeto diferente... somos capazes de perceber, incidentalmente, os processos mentais que são direcionados para esse objeto” (1874 [1973,30]). Brentano admite que a percepção interior carece necessariamente das vantagens da observação atenta, por isso recomenda juntá-la com métodos retrospectivos.

Wundt (1888) concorda com Comte e Brentano que a observação envolve necessariamente atenção e frequentemente interfere no processo a ser observado, se esse processo for interno, psicológico. Em uma extensão muito maior do que Brentano, no entanto, Wundt enfatiza a importância para a psicologia científica da atenção direta à experiência, incluindo a

variação planejada e controlada. O método psicológico de “percepção interna” é, para Wundt, o método de manter e manipular atentamente uma imagem da memória ou reprodução de um processo psicológico passado. Embora Wundt veja algum valor nesse método retrospectivo, ele acha que ele tem duas insuficiências cruciais: primeiro, só se pode trabalhar com o que se lembra do processo em questão - a manipulação de uma imagem da memória não pode descobrir novos elementos. E, segundo, elementos estranhos podem ser introduzidos involuntariamente por meio de associação - pode-se confundir a memória de um processo com a memória de outro processo ou objeto associado.

Portanto, sugere Wundt, a ciência da psicologia deve depender da observação atenta dos processos mentais à medida que eles ocorrem. Ele argumenta que aqueles que pensam que a atenção necessariamente distorce o processo mental alvo são muito pessimistas. Uma *subclasse* de processos mentais permanece relativamente imperturbável pela observação atenta - os processos mentais “mais simples”, especialmente da percepção (1896/1902, 27-28). A experiência de ver vermelho, afirma Wundt, é mais ou menos a mesma independentemente de a pessoa estar ou não atenta ao fato psicológico de estar

experienciando a vermelhidão. Wundt também sugere que os processos básicos de memória, sentimento e volição podem ser observados sistematicamente e sem perturbação excessiva. *Somente* esses, ele pensa, podem ser estudados pela psicologia introspectiva (ver também Wundt, 1874 [1904]; 1896 [1902]; 1907). Outros aspectos de nossa psicologia devem ser abordados por meio de métodos não-introspectivos, como a observação da linguagem, mitologia, cultura e desenvolvimento humano e animal.

3.3 O declínio da introspecção científica

Embora psicólogos introspectivos fossem capazes de criar um consenso científico sobre algumas questões relativas à experiência sensorial - questões como os limites da percepção sensorial em várias modalidades e alguns dos contornos da variação em experiência sensorial - no início do século 20 estava ficando claro que em muitas questões o consenso era elusivo. A disputa mais famosa dizia respeito à existência de “pensamento sem imagem” (veja a discussão da controvérsia de pensamento sem imagem no verbete representação mental⁸ [*mental imagery*];

⁸ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/mental-imagery/>>

veja também Humphrey, 1951; Kusch, 1999); mas outros tópicos se mostraram igualmente resistentes, tais como a estrutura da emoção ou “sentimento” (James, 1890 [1981]; Külpe, 1893 [1895]; Wundt, 1896 [1902]; Titchener, 1908 [1973]) e as mudanças experimentais provocadas pelas variações na atenção (Wundt, 1896 [1902]; Pillsbury, 1908; Titchener, 1908 [1973]; Chapman, 1933).

Na década de 1910, o behaviorismo (que focava unicamente na relação entre estímulo externo e resposta comportamental) havia declarado guerra à psicologia introspectiva, retratando-a como atolada em disputas irresolúveis entre distintos “especialistas” introspectivos, e repreendendo a taxonomização passiva da experiência dos introspeccionistas, recomendando que a psicologia focalizasse, em vez disso, em paradigmas socialmente úteis para modificar o comportamento (por exemplo. Watson, 1913). Nas décadas de 1920 e 1930, os estudos introspectivos foram gradativamente marginalizados. Embora o behaviorismo estrito tenha decaído nas décadas de 1960 e 1970, seu principal substituto, o funcionalismo cognitivista (que trata processos cognitivos internos definidos funcionalmente como centrais para a investigação psicológica), continuou, de modo geral, a compartilhar o desdém do behaviorismo pelos métodos introspectivos.

A psicofísica (o estudo da relação entre input sensorial física e estado ou a resposta psicológicas consequentes), na qual os psicólogos introspeccionistas haviam encontrado seu maior sucesso, passou por uma sutil mudança neste período de um foco em métodos *subjetivos* - métodos que envolvem pedir aos sujeitos para relatar suas experiências ou percepções - para um foco em métodos *objetivos*, como pedir aos participantes para relatar estados do mundo exterior, incluindo insistir que os participantes suponham mesmo quando eles sentem que não sabem ou não têm nenhuma experiência consciente relevante (especialmente com o surgimento da “teoria da detecção de sinais” na psicofísica: Green e Swets, 1966; Cheesman e Merikle, 1986; Macmillan e Creelman, 1991; Merikle, Smilek e Eastwood, 2001). Talvez, em acordo com as visões de transparência da introspecção (seção 2.3.4, acima), os dois tipos de instruções pareçam muito similares (compare o subjetivo “diga-me se você visualmente experiencia um flash de luz” com o objetivo “diga-me se a luz pisca”). Em contrapartida, talvez em tensão com visões de transparência, instruções subjetivas e objetivas parecem, às vezes, diferir significativamente, especialmente em casos de ilusão conhecida, efeitos Gestalt, como o agrupamento percebido, estímulos próximos aos limites da perceptibilidade, e a experiência de

figuras ambíguas (Boring ,1921; Merikle, Smilek, e Eastwood, 2001; Siewert, 2004).

3.4 A reemergência da introspecção científica

Em nenhum período, entretanto, os métodos introspectivos foram inteiramente abandonados pelos psicólogos e, nas últimas décadas, eles fizeram uma espécie de retorno, especialmente com o surgimento do campo interdisciplinar dos “estudos da consciência” (veja, por exemplo, Jack e Roepstorff, eds., 2003, 2004). Ericsson e Simon (1984/1993; a ser discutido na seção 4.2.3, abaixo) defenderam o uso de “protocolos pense-em-voz-alta” e de relatos imediatamente retrospectivos no estudo da solução de problemas. Outros pesquisadores enfatizaram métodos introspectivos no estudo representação [*imagery*] (Marks 1985; Kosslyn, Reisbert, e Behrmann 2006) e da emoção (Lambie e Marcel, 2002; Barrett *et al.*, 2007).

As metodologias *Beeper*⁹ [envolvendo sinalizadores sonoros aleatórios] foram desenvolvidas para facilitar a

⁹ A proposta envolve a utilização de sinalizadores sonoros por participantes de pesquisa em seus ambientes naturais. Ao ouvir os sinais sonoros emitidos aleatoriamente, através de fones de ouvido, os participantes deveriam prestar atenção a sua experiência interna em ocorrência no momento do sinal e registrar imediatamente suas características. Após um determinado número de experiências registradas, os participantes passariam por uma entrevista com os

retrospecção imediata, especialmente por Hurlburt (1990, 2011; Hurlburt e Heavey, 2006; Hurlburt e Schwitzgebel, 2007) e Csikszentmihalyi (Larson e Csikszentmihalyi, 1983; Csikszentmihalyi, 2014). Os métodos imediatamente retrospectivos tradicionais requeriam que o observador introspectivo no laboratório, de alguma forma, se abstinhasse intencionalmente de introspeccionar a experiência alvo enquanto ela ocorresse, o que é indiscutivelmente uma tarefa difícil. Hurlburt e Csikszentmihalyi, em contraste, dão aos participantes *beepers* para usarem durante atividades comuns, cotidianas. Os *beepers* são programados para soarem apenas em intervalos longos, surpreendendo os participantes e desencadeando uma avaliação imediatamente retrospectiva de sua “experiência

pesquisadores para ajudá-los a fornecer descrições as mais fiéis possíveis de suas experiências internas. O terceiro passo da metodologia prevê a elaboração de uma descrição escrita da experiência em curso em cada momento registrado. Acredita-se que o registro em momentos aleatórios permitiria alcançar a experiência original de cada pessoa, não perturbada pelo planejamento intencional da observação ou por interpretações, assegurando observações precisas e registros confiáveis. Após críticas e refinamentos ao longo dos anos, a metodologia ganhou procedimentos específicos, configurando, por exemplo, a chamada Amostragem Descritiva da Experiência (Descriptive Experience Sampling - DES, no original), contando com um número fixo de amostras de observações diárias (geralmente 6), coletadas ao longo de uma série de 4 a 8 dias de amostragem. Em um intervalo de 24h após cada registro de amostras, seria realizada a entrevista com o investigador e, na sequência, a escrita da descrição da experiência interna. Para mais detalhes, ver Hurlburt (1993, 1998), Hurlburt & Akhter (2006), Hurlburt & Heavey (2006).

interna”, emoção ou pensamentos deles no momento anterior ao *beep*.

Relatos introspectivos ou subjetivos da experiência consciente também desempenharam um papel importante na busca dos “correlatos neurais da consciência” (como revisto em Rees e Frith, 2007; Prinz, 2012; Koch *et al.*, 2016; veja também Varela, 1996). Um paradigma para os pesquisadores é apresentar estímulos sensoriais ambíguos, mantendo-os constantes durante um período prolongado, observando quais mudanças neurais se correlacionam com as mudanças nos relatos subjetivos da experiência. Por exemplo, nos métodos de “rivalidade binocular”, duas imagens diferentes (por exemplo, um rosto e uma casa) são apresentadas, uma a cada olho simultaneamente. Os participantes normalmente dizem que apenas uma imagem é visível por vez, com a imagem visível mudando a cada poucos segundos. Os pesquisadores às vezes relatam encontrar evidências de que a atividade em áreas visuais “primárias” (como V1) não está temporalmente associada às mudanças relatadas na experiência visual, enquanto as mudanças na percepção consciente estão mais adequadamente associadas temporalmente a atividade em áreas parietais e, talvez, também em áreas frontais mais próximas e a mudanças em larga escala na sincronização ou oscilação neural; contudo, a evidência é disputada (Lumer,

Friston, e Rees, 1998; Tong *et al.*, 1998; Tononi *et al.*, 1998; Polonsky *et al.*, 2000; Tong, Meng, e Blake, 2006; Kamphuisen, Bauer, e van Ee, 2008; Sandberg *et al.*, 2013; Frässle *et al.*, 2014; Ishiku e Zeki, 2014; Tsuchiya *et al.*, 2015). Outra versão do paradigma dos estímulos sensoriais ambíguos envolve apresentar ao participante uma figura ambígua, como a figura face-vaso de Rubin:



Usando esse paradigma, pesquisadores encontraram mudanças neuronais tanto em áreas visuais primárias quanto em áreas posteriores, além de mudanças na sincronia neuronal

difusa, que correspondem temporalmente a relatos subjetivos de inversões entre uma e outra forma de ver a figura ambígua (Kleinschmidt *et al.* 1998; Rodriguez *et al.* 1999; Ilg *et al.* 2008; Parkkonen *et al.* 2008; de Graaf *et al.* 2011; Megumi, Bahrami, Kanai, e Rees 2015; Brascamp, Sterzer, Blake, e Knapen 2017). Nos paradigmas de mascaramento, os estímulos são brevemente apresentados e então seguidos por uma “máscara”. Em alguns testes, os participantes relataram ter visto os estímulos, enquanto em outros não. Em testes nos quais os participantes relataram que o estímulo foi visualmente experienciado, os pesquisadores tenderam a encontrar níveis mais altos de atividade através de, pelo menos, algumas das vias visuais próximas, bem como oscilações elétricas espontâneas próximas a 40 Hz (Dehaene *et al.*, 2001; Summerfield, Jack, e Burgess, 2002; Del Cul, Baillet, e Dehaene, 2007; Quiroga *et al.*, 2008; Salti *et al.*, 2015). No entanto, permanece controverso como interpretar adequadamente tais tentativas de encontrar correlatos neurais da consciência (Noë e Thompson, 2004; Overgaard, Sandberg, e Jensen, 2008; Dehaene e Changeux, 2011; Aru, Bachmann, Singer, e Melloni, 2012; de Graaf, Hsieh, e Sack, 2012; Koch *et al.*, 2016; Phillips, 2018).

Se relatarmos nossas atitudes através da sua introspecção, então muitas das pesquisas de levantamento também são

introspectivas, embora os psicólogos geralmente não as tenham descrito explicitamente como tal. Assim como no caso dos métodos *subjetivos versus* objetivos na psicofísica, parece haver apenas uma pequena diferença entre questões subjetivamente formuladas (“Você aprova o modo como o Presidente está lidando com a guerra?”, “Você acha que a maconha deveria ser legalizada?”) e questões objetivamente formuladas (“O Presidente lidou bem com a guerra?”, “A maconha deveria ser legalizada?”). Isso parece reforçar a observação na base da teoria de transparência (discutida na seção 2.3.4, acima) de que perguntas sobre a mente e perguntas sobre o mundo exterior frequentemente exigem o mesmo tipo de reflexão.

4. A Precisão da Introspecção

4.1 Variedades de privilégio

É plausível supor que as pessoas tenham algum tipo de *acesso privilegiado* a, pelo menos, alguns de seus próprios estados ou processos mentais: você sabe sobre sua própria mente, ou pelo menos alguns aspectos dela, de um modo diferente e melhor do que você sabe sobre a mente de outras pessoas, e, talvez também de um modo diferente e melhor do que você sabe sobre o mundo exterior. Considere a dor. Parece que você conhece suas próprias dores de forma diferente e melhor do que você conhece as minhas, diferentemente e (talvez) melhor do que você conhece sobre a xícara de café em sua mão. Se assim for, talvez esse conhecimento privilegiado especial de “primeira pessoa” surja através de algo como a introspecção, em um ou mais dos sentidos descritos na seção 2, acima.

Assim como há uma diversidade de métodos para adquirir conhecimento ou alcançar juízos sobre os próprios estados e processos mentais, aos quais o rótulo “introspecção” se aplica com maior, menor ou discutível precisão, também há uma diversidade de formas de “acesso privilegiado”, com diferentes tipos de privilégios e aos quais a ideia de acesso se aplica com

maior, menor ou discutível precisão. E como seria de se esperar, os diferentes métodos introspectivos não se alinham todos igualmente bem com as diferentes variedades de privilégio.

4.1.1 Variedades de perfeição: infalibilidade, indubitabilidade, incorrigibilidade e autorrevelação

Uma tradição filosófica substancial, que remonta, pelo menos, a Descartes (1637/1985; 1641/1984; também a Augustine c. 420 C.E. /1998), atribui uma espécie de perfeição epistêmica a pelo menos alguns de nossos julgamentos (ou pensamentos, crenças ou conhecimentos) sobre nossas próprias mentes - infalibilidade, indubitabilidade, incorrigibilidade ou autorrevelação. Considere o juízo (pensamento, crença etc.) *P*, onde *P* é uma proposição que se autoatribui um estado ou processo mental (por exemplo, *P* pode ser *eu estou com dor*, ou *acredito que esteja nevando*, ou *estou pensando em um cão basset*). O juízo *P* é *infalível* apenas no caso de, se eu fizer esse juízo, não ser possível que *P* seja falso. É *indubitável* apenas no caso de, se eu fizer o juízo, não ser possível para eu duvidar da verdade de *P*. É *incorrigível* apenas no caso de, se eu fizer o juízo, não ser possível para ninguém mais mostrar que *P* é falso. E ele é autorrevelador se não for possível que *P* seja verdadeiro sem que eu chegue ao juízo (pensamento, crença etc.) de que é

verdadeiro. Note que a direção da implicação para o último desses é o inverso dos três primeiros. Infalibilidade, indubitabilidade e incorrigibilidade, todas têm a forma: “Se eu julgo (penso, acredito etc.) *P*, então ...”, enquanto a autorrevelação tem a forma: “Se *P*, então eu julgo (penso, acredito etc.) que *P*. Todas as quatro teses também admitem enfraquecimento pelo acréscimo de condições à cláusula antecedente “se” (por exemplo, “Se eu julgo que *P* como resultado de processos introspectivos normais, então ...”). (Veja Alston 1971 para uma dissecação útil dessas distinções; todas admitem variações e nuances. Note também que alguns filósofos [por exemplo., Ayer 1936/1946; Armstrong 1963; Chalmers 2003; Tye 2009] usam “incorrigibilidade” para infalibilidade, como aqui definido, enquanto outros [por exemplo, Ayer 1963; Alston 1971; Rorty 1970; Dennett 2000] usam-na com o significado mais etimologicamente específico de [algo como] “incapaz de correção”).

Descartes (1641/1984) endossou notoriamente a indubitabilidade do “eu penso”, que ele estende também a estados mentais como duvidar, entender, afirmar e parecer ter percepções sensoriais. Ele também parece reivindicar que o pensamento ou afirmação de que eu estou em tais estados é infalivelmente verdadeiro, pelo menos se esse pensamento for

claro e distinto. Ele foi seguido nisso - especialmente por seu infalibilismo - por Locke (1690 [1975]), Hume (1739 [1978]), pensadores do século XX como Husserl (1913 [1982]), Ayer (1936 [1946], 1963), Lewis (1946), no primeiro Shoemaker (1963) e muitos outros. Os argumentos históricos em favor da indubitabilidade e infalibilidade tenderam a centrar-se em apelos intuitivos à aparente impossibilidade de duvidar ou errar sobre questões como se alguém está tendo um pensamento com certo conteúdo ou se está experienciando dor ou tendo uma experiência visual, como a de ver vermelho.

Infalibilistas recentes acrescentaram a este apelo intuitivo argumentos estruturais baseados em abordagens de autorrealização de introspecção ou autoconhecimento (veja a seção 2.3.1, acima) - geralmente, ao mesmo tempo em que restringem o escopo da infalibilidade, por exemplo, a pensamentos sobre pensamentos (Burge 1988, 1996), ou a juízos fenomenais "puros" sobre a consciência (Chalmers 2003; veja também Wright 1998; Gertler 2001; Horgan, Tienson, e Graham 2006; Horgan e Kriegel 2007; Tye 2009; com importantes antecessores em Brentano 1874 [1973]; Husserl 1913 [1982]), ou a crenças como "compromissos" (Coliva 2016). A ideia intuitiva por trás da maioria destes argumentos estruturais é que, de alguma forma, o pensamento ou juízo autoatribuído *contém* o

processo ou estado mental autoatribuído: o pensamento de que estou pensando em um elefante rosa contém o pensamento de um elefante rosa; o juízo de que estou tendo uma experiência visual de vermelhidão contém a própria experiência do vermelho.

Em contraste, as abordagens de paridade eu/outro (seção 2.1) e de autodeteção (seção 2.2) da introspecção ou autoconhecimento parecem estar em tensão com o infalibilismo. Se a introspecção ou autoconhecimento envolve um processo causal de um estado mental para uma autoatribuição ontologicamente distinta desse estado, parece que, por mais confiável que tal processo possa geralmente ser, há inevitavelmente espaço, em princípio, para interferência e erro. Minimamente, parece que um acidente vascular cerebral, um acidente altamente improvável ou uma neurocirurgia inteligente poderia quebrar relações entre os estados mentais alvo e as autoatribuições desses estados, relações que, de outro modo, seriam geralmente confiáveis. Considerações similares aplicam-se as abordagens de automodelagem (seção 2.3.2) e expressivistas (seção 2.3.3), na medida em que estas são interpretadas causalmente em vez de constitutivamente.

A incorrigibilidade introspectiva, em oposição tanto à infalibilidade quanto à indubitabilidade, foi considerada por Rorty (1970) como “a marca do mental” – e, assim, como se

aplicando a uma ampla gama de estados mentais. Dennett (2000, 2002) defende uma visão similar para experiências conscientes. A ideia por trás da incorrigibilidade, ele lembra, é que ninguém mais poderia mostrar que suas autoatribuições são falsas; ou poderíamos dizer, mais qualificadamente e de forma um pouco diferente, que se você chegar ao tipo certo juízo autoatributivo (talvez um juízo introspectivamente baseado sobre um processo consciente atualmente em curso que sobrevive à reflexão crítica), então ninguém mais, talvez nem mesmo você no futuro, consciente [*aware of*] disso, pode racionalmente considerar esse juízo como estando errado. Se eu julgo que neste momento estou sentindo uma dor severa, e o faço como resultado de considerar introspectivamente se estou de fato com tal dor (em oposição a, digamos, meramente inferir que estou com dor me baseando no comportamento externo), e se eu parar para pensar cuidadosamente sobre se estou realmente com dor e concluir que de fato estou, então ninguém mais que está ciente disso pode racionalmente acreditar que eu não estou com dor, independentemente de qual possa ser meu comportamento externo (digamos, calmo e relaxado) ou do que aparece no curso da imagem cerebral (digamos, nenhuma ativação em centros cerebrais normalmente associados à dor).

Incorrigibilidade não implica em infalibilidade: eu posso não estar de fato com dor, mesmo que ninguém possa *mostrar* que não estou. Consequentemente, a incorrigibilidade é compatível com uma maior variedade de fontes de autoconhecimento do que a infalibilidade. Nem Rorty nem Dennett, por exemplo, parecem defender a incorrigibilidade apelando para abordagens de autorrealização da introspecção (embora, em ambos os casos, interpretar suas abordagens positivas seja difícil). As abordagens causais de autoconhecimento podem ser compatíveis com a incorrigibilidade se as conexões causais que garantem os juízos incorrigíveis forem muito mais confiáveis do que os juízos obtidos sem o benefício deste tipo de acesso privilegiado. É claro que, a menos que se adote uma abordagem estrita de autorrealização, com o infabilismo que tal perspectiva traz, se quererá excluir casos anormais como o acidente altamente improvável; daí a necessidade de fazer qualificações.

Os estados mentais autorreveladores são aqueles que, se uma pessoa (ou, pelo menos, uma pessoa com as capacidades necessárias adequadas) os possui, ela necessariamente acredita, julga ou sabe que os possui. Os estados conscientes são frequentemente considerados, em algum sentido, autorreveladores, na medida em que o simples fato de os ter

envolve, requer ou implica algum tipo de representação ou consciência desses estados. Brentano argumenta que a consciência, por exemplo, de um estímulo externo como um som, “ocorre claramente junto com a consciência desta consciência”, isto é, a consciência é “de todo ato mental no qual o som é apresentado e no qual a própria consciência existe concomitantemente” (1874 [1995, 129]; veja também abordagens fenomenológicas para autoconsciência). Teorias da consciência de “ordem superior” e “mesma ordem” (Armstrong 1968; Rosenthal 1990, 2005; Gennaro 1996; Lycan 1996; Carruthers 2005; Kriegel 2009; veja também teorias da consciência de ordem superior) explicam a consciência em termos de determinado pensamento, percepção ou representação do estado mental que é consciente – sendo a presença desse pensamento, percepção ou representação sendo o que torna o estado alvo consciente. (Nas teorias de mesma ordem, o estado mental alvo, ou um aspecto dele, representa a si mesmo, sem necessidade de um estado de ordem superior distinto). Assim, Horgan, Kriegel, e outros descreveram a consciência como “autoapresentação”. (Horgan, Tienson, e Graham 2005; Horgan e Kriegel 2007; o uso parece seguir Chisholm 1981, mas Chisholm, na verdade, tem em mente uma tese de indubitabilidade em vez de uma tese de autorrevelação). Shoemaker (1995, 2012) argumenta que crenças

- desde que estejam “disponíveis” (ou seja, prontamente implantadas em inferência, consentimento, raciocínio prático etc.), não precisam requerer que elas sejam conscientes no momento – são autorreveladoras para indivíduos com capacidade cognitiva suficiente. A ideia de Shoemaker é que se a crença de que *P* está disponível no sentido relevante, então se está disposto a fazer coisas como dizer “eu acredito em *P*”, e tais disposições são, elas próprias, constitutivas do acreditar que se acredita que *P* [seja o caso].

Reivindicações de autorrevelação (ao contrário das reivindicações de infalibilidade, indubitabilidade e incorrigibilidade) geralmente não são classificadas como reivindicações sobre “introspecção”. Isso pode ser porque o conhecimento adquirido através da autorrevelação pareceria ser constante e automático, violando assim a condição de esforço na introspecção (condição 6 na seção 1.1, acima).

4.1.2 Garantias mais fracas

Vários filósofos defendem formas de privilégio em primeira pessoa que envolvem algum tipo de garantia epistêmica - não apenas precisão condicional como uma questão de fato empírico, mas algo mais robusto do que isso - sem abranger infalibilidade,

indubitabilidade, incorrigibilidade ou autorrevelação nos sentidos descritos na Seção 4.1.1, acima.

Shoemaker (1968), por exemplo, argumenta que o autoconhecimento de certos fatos psicológicos como “estou balançando meu braço” ou “eu vejo um canário”, quando alcançado “da maneira ordinária (sem o auxílio de espelhos, etc.)”, *é imune ao erro por identificação errônea relativa ao pronome de primeira pessoa* (veja também Campbell 1999; Pryor 1999; Bar-On 2004; Hamilton 2008). Ou seja, embora seja possível estar errado sobre acenar com o braço (talvez os nervos no seu braço tenham sido recentemente rompidos sem o seu conhecimento) ou sobre ver um canário (talvez seja um pintassilgo), não se pode estar errado por identificar incorretamente a pessoa acenando o braço ou vendo o canário *como você*, quando na verdade é outra pessoa. Essa imunidade surge, argumenta Shoemaker, porque não há necessidade de identificação em primeiro lugar e, portanto, não há oportunidade para erros de identificação. A este respeito, Shoemaker argumenta, o conhecimento de que um determinado braço que está se movendo é o seu braço (não imune ao erro de identificação já que talvez seja o braço de outra pessoa erroneamente identificada no espelho) é diferente do conhecimento de que você está movendo seu braço -

conhecimento, aliás, do que Searle (1983) chama de "intenção em ação".

Shoemaker também defendeu a impossibilidade conceitual da autocegueira introspectiva no que diz respeito às crenças, desejos e intenções que se tem e, por razões um tanto diferentes, às dores que se tem (1988, 1994b). Uma criatura cega para si mesma, pela definição de Shoemaker, seria uma criatura racional com uma concepção dos estados mentais relevantes, e que pode considerar o pensamento de que ela tem essa ou aquela crença, desejo, intenção ou dor, mas que, no entanto, carece totalmente de acesso introspectivo ao tipo de estado mental em questão. Uma criatura cega para si mesma ainda poderia obter conhecimento de “terceira pessoa” dos estados mentais em questão, por meio da observação do seu próprio comportamento, lendo livros didáticos e semelhantes (Por isso, abordagens do autoconhecimento de paridade eu/outro estritas, do tipo descrito na Seção 2.1, são abordagens segundo as quais é cego para si mesmo no sentido de Shoemaker.) O caso de Shoemaker contra a autocegueira no que diz respeito à crença se volta para o dilema de se a criatura cega para si mesma pode evitar as sentenças “paradoxais de Moore” (veja Moore 1942, 1944 [1993]; Shoemaker 1995) como “chove, mas não creio que chove” em que o sujeito afirma tanto *P* quanto afirma que não acredita que

P [seja o caso]. Se o sujeito é verdadeiramente cego para si mesmo, sugere Shoemaker, deveria haver casos em que sua melhor evidência é simultaneamente que P [é o caso] e que ele não acredita que P [seja o caso] (este último, talvez, baseado em fatos enganosos sobre seu comportamento). Mas se o sujeito afirma “ P [é o caso], mas eu não acredito que P [seja o caso]” em tais casos, ele não tem (contra a suposição inicial) realmente um comando racional da natureza da crença e da asserção; e, portanto, não é um caso genuíno de autocegueira como originalmente previsto. Alternativamente, talvez a criatura possa evitar, de forma confiável, tais sentenças paradoxais de Moore, autoatribuindo a crença de forma aparentemente normal. Mas então, sugere Shoemaker, parece que ela é indistinguível das pessoas normais em pensamento e comportamento e, portanto, não é cega para si mesma. Para o desejo, a intenção e a dor, Shoemaker pretende revelar incoerências entre ter um controle racional dos conceitos em questão e se comportar como se fosse sistematicamente ignorante ou equivocado sobre esses estados. Shoemaker usa seu caso contra a autocegueira como parte de seu argumento contra abordagens de autodetecção da introspecção (descrito na seção 2.2, acima): se a introspecção fosse uma questão de detectar a presença de estados que existem independentemente do juízo ou crença introspectiva, então

deveria ser possível para a faculdade permitir que a detecção tivesse uma ruptura total, como no caso da cegueira, surdez, etc., na percepção externa (veja também Nichols e Stich 2003, que argumentam que a esquizofrenia fornece tal caso).

Burge influentemente afirmou que *erros brutos* sobre “atitudes proposicionais presentes, comuns e acessíveis [como crença e desejo]” são impossíveis ou, pelo menos, sujeitos a “limites severos” - onde um “erro bruto” é um erro que “não indica falha racional e nem mau funcionamento no indivíduo equivocado” como comumente ocorre na percepção comum devido a “condições naturais enganosas ou substitutos parecidos” (1988, 657-658; 1996, 103-104). Entretanto, Burge oferece poucos argumentos para esta alegação, além daquele mencionado nas seções 2.3.1 e 4.1.1, acima, de que, para certos tipos de autoatribuições, o erro em geral (e não apenas o “erro bruto”) é impossível, devido à natureza “autoverificadora” de tais autoatribuições.

Dretske (1995, 2004) argumenta que temos um conhecimento infalível do *conteúdo* de nossas atitudes sem necessariamente conhecer (ou mesmo ter uma ideia muito boa sobre) a *atitude* que tomamos em relação a esses conteúdos. Por exemplo, se eu acredito que *vai chover amanhã*, eu tenho informações infalivelmente precisas, às quais eu posso então

acessar introspectivamente, relativamente à presença de um estado mental com um certo conteúdo - o conteúdo “vai chover amanhã” - mas eu posso frequentemente ter pouca ou nenhuma informação sobre o fato de que minha atitude em relação a esse conteúdo é a atitude particular que ela é - crença, nesse caso (em oposição a suposição ou esperança). Essa visão decorre da aceitação por Dretske de algo como uma abordagem de contenção sobre a introspecção do conteúdo da atitude (o juízo introspectivo empregando o mesmo conteúdo da atitude alvo; ver seção 2.3.1, acima, especialmente a discussão de Burge), enquanto ele vê o conhecimento da atitude que se tem em relação a esse conteúdo como algo que requer informações complexas sobre o papel causal e a história desse estado mental.

Argumentos transcendentais para a precisão de certos tipos de autoconhecimento oferecem um tipo diferente de garantia epistêmica - “argumentos transcendentais” sendo argumentos que assumem a existência de algum tipo de experiência ou capacidade e, então, desenvolvem *insights* sobre as condições de fundo necessárias para essa experiência ou capacidade e, finalmente, concluem que essas condições de fundo devem, de fato, ser satisfeitas. Burge (1996; ver também Shoemaker 1988) argumenta que, para ser capaz de “raciocínio crítico”, é preciso ser capaz de reconhecer as próprias atitudes, avaliar, identificar e rever, com

conhecimento de causa, as próprias crenças, desejos, comprometimentos, suposições etc., onde esses estados mentais são conhecidos por serem os estados que eles são. Uma vez que somos (por suposição, em nome do argumento transcendental) capazes de raciocínio crítico, devemos ter algum conhecimento de nossas atitudes. Bilgrami (2006) argumenta que só podemos ser responsabilizados por ações se conhecermos as crenças e desejos que “racionalizam” nossas ações; uma vez que, às vezes, podemos (por suposição) ser responsabilizados, devemos ocasionalmente conhecer nossas crenças e desejos. Wright (1989) argumenta que o “jogo de linguagem” de atribuir “estados intencionais”, tais como crença e desejo, a si mesmo e aos outros, requer como condição de fundo que as autoatribuições tenham autoridade especial dentro desse jogo. Dado que jogamos este jogo de linguagem com sucesso, devemos de fato ter a autoridade especial que assumimos ter e que outros nos concedem no contexto do jogo.

4.1.3 Privilégio sem garantia

Desenvolvendo uma analogia a partir de Wright (1998), se for a sua vez com o caleidoscópio, você tem um tipo de perspectiva privilegiada das formas e cores que ele apresenta. Se outra pessoa na sala quiser saber qual cor predomina, por

exemplo, o caminho mais direto seria perguntar a você. Mas esse tipo de acesso privilegiado não traz nenhuma garantia. Pelo menos a princípio, você pode estar bem errado em relação às formas em queda. Você pode ficar deslumbrado por imagens residuais, momentaneamente confuso, com alucinações ou daltônico (sem que você saiba). (Sim, muitas vezes as pessoas não sabem que são daltônicas, um ponto enfatizado por Kornblith 1998). Também é possível, pelo menos em princípio, que outros saibam melhor do que você, talvez até mesmo sistematicamente, o que está acontecendo no caleidoscópio. Você pode pensar que a figura mostra uma simetria octogonal, mas o resto de nós, familiarizados com o modelo do caleidoscópio, podemos saber que a simetria é hexagonal. Um engenheiro brilhante pode inventar um detector de estado de caleidoscópio que possa revelar, de forma confiável e a partir do exterior a forma, a cor e a posição dos fragmentos em queda.

Wright levanta essa analogia para sugerir que o privilégio das pessoas, no que diz respeito a certos aspectos de suas vidas mentais, deve ser diferente daquele da pessoa com o caleidoscópio; mas outros filósofos, especialmente aqueles que adotam teorias de autodeteção da introspecção, deveriam achar a analogia pelo menos adequada em certa medida: o privilégio introspectivo é similar ao privilégio de ter uma perspectiva

sensorial única e vantajosa sobre algo. Metaforicamente falando, somos os únicos que podemos olhar diretamente para nossas atitudes ou nosso fluxo de experiência, enquanto os outros devem confiar em nós ou em sinais externos. De forma menos metafórica, ao gerar juízos introspectivos (ou crenças ou conhecimento) sobre a própria mentalidade, emprega-se um processo de detecção não disponível a mais ninguém. É então uma questão empírica o quão precisas são as afirmações deste processo; mas, partindo do pressuposto de que as declarações são, em uma ampla gama de condições, pelo menos em certa medida precisas e mais precisas do que os juízos típicos que outras pessoas fazem sobre esses mesmos aspectos de sua mente, você tem uma perspectiva “privilegiada”. Normalmente, os defensores dos modelos de introspecção de autodetecção consideram o mecanismo ou processo cognitivo que gera juízos ou crenças introspectivas, grosso modo, como altamente confiável, mas não infalível e não imune à correção por outras pessoas. (Armstrong 1968; Churchland 1988; Hill 1981, 2009; Lycan 1996; Nichols e Stich 2003; Goldman 2000, 2006).

4.2 Evidência empírica sobre a precisão da introspecção

Os argumentos da seção anterior são a priori, pelo menos no sentido amplo desse termo (o sentido dos psicólogos): eles dependem de considerações conceituais gerais e da psicologia dos povos de gabinete, e não de pesquisa empírica. A estes pode ser adicionado o argumento, graças a Boghossian (1989), de que o “externalismo” sobre o conteúdo de nossas atitudes (a visão de que nossas atitudes dependem constitutivamente não apenas do que está acontecendo internamente, mas também de fatos sobre nosso ambiente; Putnam 1975; Burge 1979) parece problematizar o autoconhecimento introspectivo dessas atitudes. Essa questão não será tratada aqui, uma vez que é amplamente abordada nos verbetes acerca do Externalismo sobre conteúdo mental¹⁰ e Externalismo e autoconhecimento¹¹.

Agora nos voltamos para a pesquisa empírica sobre nosso autoconhecimento daqueles aspectos de nossas mentes frequentemente considerados acessíveis à introspecção. Uma vez que os traços de caráter não são geralmente considerados aspectos de nossa mentalidade passíveis de introspecção,

¹⁰ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/content-externalism/>>

¹¹ N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/self-knowledge-externalism/>>

pularemos a vasta literatura acerca da precisão ou imprecisão de nossos juízos sobre eles (por exemplo, Taylor e Brown 1988; Paulhus e John 1998; Funder 1999; Vazire 2010; veja também a perspectiva cética de Haybron em 2008 sobre nosso conhecimento de quão felizes somos); também não discutiremos o autoconhecimento de processos mentais subpessoais e não-conscientes, como os processos subjacentes ao reconhecimento visual de cor e forma.

De modo geral, enquanto as explicações a priori da epistemologia da introspecção tendem a enfatizar seu privilégio e precisão, as explicações empíricas tendem a enfatizar suas falhas.

4.2.1 Das causas das atitudes e comportamentos

Talvez o argumento mais famoso na literatura psicológica sobre introspecção e autoconhecimento seja o argumento de Nisbett e Wilson de que temos um conhecimento notavelmente pobre das causas e processos subjacentes aos nossos comportamentos e atitudes (Nisbett e Wilson 1977; Nisbett e Ross 1980; Wilson 2002). A seção 2.1, acima, mencionou brevemente sua descoberta emblemática de que as pessoas em um centro comercial frequentemente ignoravam um fator importante - posição - influenciando seus juízos sobre a qualidade dos pares de meias. Em Nisbett e Bellows (1977), também

mencionados brevemente acima, os participantes foram solicitados a avaliar a influência de vários fatores em seus juízos sobre as características de um suposto candidato a emprego. Como no estudo de meias de Nisbett e Wilson, os participantes negaram a influência de alguns fatores que eram de fato influentes; por exemplo, eles negaram que a informação de que iriam se encontrar com o candidato influenciasse seus juízos sobre a flexibilidade do candidato. (Isso, na verdade, teve uma grande influência, aferida ao comparar os juízos dos participantes que foram informados e não informados de que encontrariam o candidato.) Os participantes também atribuíram influência a fatores que não eram de fato influentes; por exemplo, eles relataram falsamente que a informação de que o candidato acidentalmente derrubou uma xícara de café durante a entrevista influenciou “o quão simpática a pessoa pareceu” a eles. Nisbett e Bellows descobriram que as avaliações hipotéticas dos observadores comuns sobre a influência dos vários fatores sobre vários juízos eram muito semelhantes às avaliações dos próprios participantes sobre os fatores que os influenciam - uma descoberta usada por Nisbett para argumentar que as pessoas não têm acesso especial a influências causais em seus juízos e, em vez disso, confiam nos mesmos tipos de considerações teóricas em que os observadores externos confiam (a visão de paridade

eu/outro é descrita na Seção 2.1). Apesar de algumas objeções (como White 1988), tanto psicólogos quanto filósofos agora tendem a aceitar a visão de Nisbett e Wilson de que há, na melhor das hipóteses, apenas uma vantagem modesta da posição de primeira pessoa na avaliação dos fatores que influenciam nossos juízos e comportamento.

Em uma série de experimentos, Gazzaniga (1995) apresentou aos pacientes submetidos à comissurotomia (pessoas que tiveram o corpo caloso seccionado) estímulos visuais diferentes para cada hemisfério do cérebro. Com a comunicação hemisférica gravemente prejudicada devido à comissurotomia, o hemisfério esquerdo, controlando a fala, tinha informações sobre uma parte do estímulo visual, enquanto o hemisfério direito, controlando alguns aspectos do movimento (especialmente a mão esquerda), tinha informações sobre uma parte diferente. Gazzaniga relatou que, quando esses pacientes de “cérebro dividido” eram solicitados a explicar por que fizeram algo, quando essa ação era claramente causada por um *input* no hemisfério direito, não-verbal, o hemisfério esquerdo às vezes confabulava fluentemente uma explicação. Por exemplo, Gazzaniga relata apresentar uma instrução como “rir” para o hemisfério direito, fazendo o paciente rir. Quando perguntado por que ria, o paciente dizia algo como “Vocês vêm e nos testam todo mês. Que

maneira de ganhar a vida!” (1893). Quando uma garra de galinha era mostrada no hemisfério esquerdo e uma cena de neve ao direito, e ao paciente era solicitado a selecionar de um conjunto uma imagem apropriada, a mão direita apontava para uma galinha e a mão esquerda para uma pá de neve, e quando perguntado por que eles selecionaram essas duas coisas, o paciente diria algo como “Ah, isso é simples. A garra de galinha vai com a galinha e você precisa de uma pá para limpar o galinheiro” (*ibid.*). Confabulações semelhantes sobre motivos são às vezes (mas nem sempre) vistas em pessoas cujo comportamento é, sem que elas saibam, impulsionado por sugestão pós-hipnótica (Richet 1884; Moll 1889[1911]), e em distúrbios como heminegligência (anosognosia), negação da cegueira (síndrome de Anton) e síndrome de Korsakoff (Hirstein 2005).

Em uma população normal, Johansson e colaboradores (Johansson *et al.* 2005; Johansson *et. al.* 2006) exibiram manualmente aos participantes pares de fotos de rostos de mulheres. Em cada prova, o participante deveria apontar para o rosto que achava mais atraente. A imagem desse rosto foi então centralizada diante do participante enquanto o outro rosto estava escondido. Em algumas provas, aos participantes foi solicitado explicar as razões de suas escolhas enquanto continuavam a olhar

para o rosto selecionado. Em algumas provas importantes, os experimentadores usaram truques de ilusionismo para apresentar ao participante o rosto que *não* foi selecionado como se tivesse sido o rosto selecionado. Surpreendentemente, a mudança foi notada apenas 28% das vezes. Além disso, quando a mudança não foi detectada, os participantes realmente deram explicações para sua escolha, que apelaram para características específicas do rosto não selecionado que não eram possuídas pelo rosto selecionado em 13% das vezes. Por exemplo, um participante afirmou ter escolhido o rosto diante dele “porque eu amo loiras”, quando na verdade ele escolheu um rosto de cabelos escuros (Johansson et. al 2006, 690). Johansson e colegas não conseguiram encontrar quaisquer diferenças sistemáticas nas explicações de escolha entre os ensaios manipulados e os não manipulados, usando uma ampla variedade de medidas. Eles não encontraram, por exemplo, nenhuma diferença nos marcadores linguísticos de confiança (incluindo pausas na fala), emotividade, especificidade de detalhes, complexidade ou extensão da descrição ou posição geral no espaço semântico. Esses resultados, como os de Nisbett e Wilson, sugerem que pelo menos parte do tempo em que as pessoas pensam que estão explicando as bases de suas decisões, elas estão meramente teorizando ou confabulando.

Wegner descobriu que as pessoas podem frequentemente ser manipuladas a acreditar que desejaram ou pretenderam realizar um comportamento que é de fato causado pela manipulação de outra pessoa e, inversamente, que elas não exercem controle sobre movimentos que eram de fato seus próprios – como nos tabuleiros *ouija*, com ou sem um cúmplice trapaceiro e intencionalmente diretivo (Wegner e Wheatley 1999; Wegner 2002). A literatura sobre “dissonância cognitiva” está repleta de casos em que as atitudes dos participantes parecem mudar por razões que eles negam ou negariam. De acordo com a teoria da dissonância cognitiva, quando as pessoas se comportam ou parecem se comportar de forma contranormativa (por exemplo, de forma incompleta, tola ou imoral), elas tendem a ajustar suas atitudes de modo a tornar o comportamento menos contranormativo ou “dissonante” (Festinger 1957; Aronson 1968; Cooper e Fazio 1984; Stone e Cooper 2001). Por exemplo, pessoas induzidas a descrever falsamente como agradável uma tarefa monótona que acabaram de concluir tenderão, mais tarde, a relatar ter uma atitude mais positiva em relação à tarefa do que aquelas que não foram induzidas a mentir (embora muito menos frequentemente se fossem generosamente pagas para mentir, caso em que o comportamento não é claramente contra normativo; Festinger e Carlsmith 1959; mas veja Bem 1967, 1972

para um argumento de que a atitude não muda, apenas o relato dela). Presumivelmente, se tais mudanças de atitude fossem conhecidas pela pessoa, elas geralmente não teriam seu efeito de redução de dissonância. Psicólogos pesquisadores também confirmaram fenômenos familiares como “uvas azedas” (Elster 1983/2016; Lyubomirsky e Ross 1999; Kay, Jiminez, e Jost 2002) e “autoengano” (Mele 2001) que presumivelmente também envolvem ignorância dos fatores conduzindo os juízos e ações relevantes. E, é claro, a tradição psicanalítica freudiana também sustenta há muito tempo que as pessoas frequentemente têm apenas um conhecimento insuficiente de seus motivos e das influências em suas atitudes (Wollheim 1981; Cavell 2006).

À luz dessa pesquisa empírica, nenhum grande filósofo agora sustenta (talvez nenhum grande filósofo jamais tenha sustentado) que temos conhecimento infalível, indubitável, incorrigível ou autorrevelador das causas de nossos juízos, decisões e comportamento. Talvez formas mais fracas de privilégio também estejam ameaçadas. Mas surge a pergunta: quaisquer que sejam as falhas que possam existir na avaliação das causas de nossas atitudes e comportamento, são elas falhas de *introspecção*, devidamente interpretadas? Os psicólogos tendem a considerar esses resultados como falhas de “introspecção”, mas, se for descoberto que um processo muito diferente e mais

confiável garante nosso conhecimento de alguns outros aspectos de nossas mentes - tais como quais são nossas atitudes atuais (independentemente de como sejam causadas) ou nossa experiência consciente atualmente em andamento ou ocorrida recentemente -, então talvez possamos chamar apenas aquele processo de introspecção, retendo assim alguma forma robusta de privilégio introspectivo enquanto aderimos ao consenso psicológico em reação ao conhecimento de causas em primeira pessoa (o que hoje chamaríamos de não introspectivo). De fato, poucas abordagens filosóficas contemporâneas de introspecção ou autoconhecimento privilegiado destacam como o *locus* primário de privilégio as causas de nossas atitudes e comportamento (embora Bilgrami 2006 seja uma exceção notável). Assim, a literatura revisada nesta seção pode ser interpretada como sugerindo que as causas de nosso comportamento não são, afinal, os tipos de coisas às quais temos acesso introspectivo.

4.2.2 De atitudes

Psicólogos pesquisadores geralmente não têm sido tão céticos em relação ao nosso conhecimento sobre nossas atitudes quanto têm sido sobre o nosso conhecimento das causas de nossas atitudes (Seção 4.2.1, acima). De fato, muitos dos mesmos

experimentos que pretendem mostrar um conhecimento impreciso das causas de nossas atitudes ainda assim dependem, de forma incauta, do autorrelato para avaliação das atitudes em si - uma característica daqueles experimentos criticados por Bem (1967). Pesquisas de levantamento sobre atitude na psicologia e nas ciências sociais frequentemente se baseiam no autorrelato dos participantes como a principal fonte de evidência sobre atitudes (de Vaus 1985/2022; Sirken *et al.* (eds.) 1999). Contudo, como no caso dos motivos e causas, há uma longa tradição na psicologia clínica de ceticismo quanto ao nosso autoconhecimento de nossas atitudes, dando um grande papel aos motivos e atitudes “inconscientes”.

Um desafio fundamental na avaliação da precisão das crenças ou juízos das pessoas sobre suas atitudes é a dificuldade de medir com precisão as atitudes independentemente do autorrelato. Atualmente, não há nenhuma medida tratável de atitude que seja vista pelos filósofos em geral como prevalecendo sobre os próprios relatos dos indivíduos sobre suas atitudes. Entretanto, na literatura psicológica, medidas “implícitas” das atitudes - medidas de atitudes que não correspondem ao autorrelato - vem recentemente ganhando considerável atenção (ver Wittenbrink e Schwarz, eds., 2007; Petty, Fazio e Briñol ed., 2009). Tais medidas são às vezes consideradas capazes de revelar

atitudes inconscientes ou implícitas, indisponíveis à introspecção ou erroneamente introspectadas (Wilson, Lindsey e Schooler 2000; Kihlstrom 2004; Lane *et al.* 2007; consulte Hahn *et al.* 2014).

Muito da pesquisa mais importante sobre medidas implícitas das atitudes se ocupou do racismo, de acordo com a visão de que atitudes racistas, embora comuns, são consideradas socialmente indesejáveis e, portanto, frequentemente não são autorreconhecidas, mesmo quando presentes. Por exemplo, Campbell, Kruskal e Wallace (1966) exploraram o uso da distância entre os assentos como um índice de atitudes raciais, observando que os alunos racialmente negros e brancos tendiam a se agregar nos arranjos de assentos em sala de aula. Usando eletromiografia facial (EMG), Vanman e outros (1997) descobriram que participantes (racialmente) brancos exibiam respostas faciais indicativas de afeto negativo com mais frequência quando solicitados a imaginar atividade cooperativa com parceiros negros do que com parceiros brancos - resultados interpretados como indicativos de atitudes racistas. Cunningham *et al.* (2004) mostraram rostos brancos e negros para participantes brancos enquanto os participantes estavam sendo submetidos a exames de imagens cerebrais de fMRI. Eles encontraram menos ativação da amígdala quando os participantes olharam para rostos

de seu próprio grupo do que quando os participantes olharam para outros rostos; e como a ativação da amígdala é geralmente associada à emoção negativa, eles interpretaram essa tendência sugerindo uma atitude negativa em relação aos membros do outro grupo (ver também Hart *et al.* 1990; e a discussão Ito e Cacioppo 2007).

Muito da pesquisa recente sobre atitudes implícitas se concentrou no *priming* de respostas e na interferência em tarefas aceleradas. Na pesquisa de *priming*, um estímulo (o “*prime*”) é exibido brevemente, seguido por uma máscara que o oculta e, em seguida, um segundo estímulo (o “alvo”) é exibido. A tarefa do participante é responder o mais rápido possível ao alvo, normalmente com um juízo de classificação. No *priming avaliativo*, por exemplo, o participante é preparado com uma palavra ou imagem com valência positiva ou negativa (por exemplo, cobra), e então solicitado a fazer um rápido juízo sobre se a palavra alvo apresentada posteriormente (por exemplo, “nojento”) é boa ou ruim, ou tem alguma outra característica (por exemplo, pertence a uma categoria particular). Geralmente, os *primes* negativos acelerarão a resposta para alvos negativos enquanto atrasam a resposta para alvos positivos, e os *primes* positivos farão o inverso. Pesquisadores descobriram que fotografias de rostos negros, sejam apresentadas visivelmente ou

tão rapidamente a ponto de serem subliminares, tendem a facilitar a categorização de alvos negativos e atrasar a categorização de alvos positivos para participantes brancos - um resultado amplamente interpretado como revelador de atitudes racistas (Fazio *et al.* 1995; Dovidio *et al.* 1997; Wittenbrink, Judd e Park 1997). No *Teste de Associação Implícita*, os entrevistados são solicitados a responder disjuntivamente a categorias combinadas, dando, por exemplo, uma resposta se virem um rosto de pele escura ou uma palavra de valência positiva e uma resposta diferente se virem um rosto de pele clara ou uma palavra de valência negativa. Como nas tarefas de *priming* avaliativo, os entrevistados brancos tendem a responder mais lentamente quando solicitados a emparelhar rostos de pele escura com palavras de valência positiva do que com palavras de valência negativa, o que é interpretado como revelando uma atitude ou associação negativa (Greenwald, McGhee e Schwartz 1998; Lane *et al.* 2007). No entanto, deve-se notar que, apesar de sua proeminência, a Tarefa de Associação Implícita foi recentemente criticada por ter baixa confiabilidade de teste-reteste e fracas correlações com outras medidas de racismo (Oswald, Mitchell, Blanton, Jaccard e Tetlock 2013; Gawronski, Morrison, Phills e Galdi 2017; Payne, Vuletich e Lundberg 2017; Signal 2017).

Como mencionado acima, tais medidas implícitas são frequentemente interpretadas como reveladoras de atitudes às quais as pessoas têm pouco ou nenhum acesso introspectivo. A evidência de que as pessoas carecem de conhecimento introspectivo de tais atitudes geralmente se baseia nas baixas correlações entre tais medidas implícitas de racismo e medidas mais explícitas, como o autorrelato - embora, devido à reconhecida indesejabilidade social do preconceito racial, seja difícil separar os fatores de autoapresentação de fatores de autoconhecimento em autorrelatos (Fazio *et al.* 1995; Greenwald, McGhee e Schwartz 1998; Wilson, Lindsey e Schooler 2000; Greenwald e Nosek 2009). Pessoas que parecem racistas por meio de medidas implícitas podem repudiar o racismo e inibir padrões racistas de resposta em medidas explícitas (como quando solicitadas a avaliar a atratividade de rostos de diferentes raças) porque não querem ser *vistas* como racistas - uma motivação que pode conduzi-las, independentemente de terem ou não um autoconhecimento preciso de suas atitudes racistas. Ainda assim, parece plausível à primeira vista que as pessoas tenham, no melhor dos casos, conhecimento limitado dos padrões de associação que conduzem suas respostas sob o *priming* e outras medidas implícitas.

Mas o que esses testes realmente medem? Em filosofia, Zimmerman (2018) e Gendler (2008a, 2008b) argumentam que medidas como o Teste de Associação Implícita não medem crenças racistas reais, mas sim outra coisa, algo sob controle menos racional (Gendler as chama de “*aliefs*”). Schwitzgebel (2010) argumenta que as pessoas que são implicitamente preconceituosas, mas explicitamente igualitárias, estão “no meio” entre acreditar e não acreditar nas proposições igualitárias que aceitam sinceramente (ver também Levy 2015). Machery (2016) argumenta que as medidas implícitas revelam disposições multifacetadas, em vez de atitudes. Gawronski e Bodenhausen (2006) propõem um modelo segundo o qual há uma diferença substancial entre atitudes implícitas, definidas em termos de processos associativos, e atitudes explícitas, que têm uma estrutura proposicional e são guiadas por padrões de verdade e consistência (ver também Wilson, Lindsey e Schooler 2000; Freenwald e Nosek 2009). Mandelbaum (2016) e Borgoni (2016) também endossam visões de que pessoas implicitamente preconceituosas, mas explicitamente igualitárias, têm atitudes contraditórias, embora argumentem que tanto as atitudes implícitas quanto as explícitas são estruturadas proposicionalmente e, em certa medida, sujeitas a normas de racionalidade. Payne, Vuletich e Lundberg (2017) argumentam

que as medidas implícitas capturam a acessibilidade situacionalmente variável de conceitos culturalmente dados.

A forma como se responde a esta pergunta sobre a relação entre viés implícito e crença ou outras atitudes têm a ver com a questão da precisão da introspecção da crença ou outra atitude em questão. Sob a suposição de que as pessoas não estão conscientes [*unaware*] da extensão de seu viés, ou pelo menos não têm acesso introspectivo direto ao seu viés, aquela falha da introspecção do viés constitui uma falha da introspecção em relação à atitude em questão. Por outro lado, se o que está em jogo é meramente uma associação ou uma disposição semelhante a um traço, em vez de uma atitude, o fracasso da introspectabilidade não é surpreendente e não tem relação com a questão geral da introspecção das atitudes.

A questão se generaliza além do viés implícito. Na medida em que as atitudes são consideradas refletidas em, ou mesmo definidas por nossos juízos explícitos sobre o assunto em questão e, de forma diferente, mas talvez não totalmente separável (ver Seção 2.3.4, acima), por nossos juízos explícitos sobre nossas *atitudes* em relação ao assunto em questão, nosso autoconhecimento pareceria ser correspondentemente seguro o implícito é medido para além do ponto. Na medida em que se considera que as atitudes envolvem, de forma crucial, padrões de

reação e associação rápidos e automáticos, ou irrefletidos, nosso autoconhecimento deles parece ser correspondentemente problemático, corrigível por dados de medidas implícitas (Bohner e Dickel 2011; Schwitzgebel 2011a, no prelo).

Da mesma forma, Carruthers (2011; ver também Bem 1967, 1972; Rosenthal 2001; Cassam 2014) argumenta que a evidência de Nisbett, Gazzaniga, Wegner e outros (revisadas na Seção 4.2.1, acima) mostram que as pessoas confabulam não apenas ao relatar as causas de suas atitudes, mas também ao relatar as próprias atitudes. Por exemplo, Carruthers sugere que se alguém no famoso estudo de Nisbett e Wilson de 1977 confabula “Eu pensei que este par era o mais macio” como uma explicação de sua escolha do par de meias mais à direita, ele erra não apenas sobre a causa de sua escolha, mas ao autoatribuir o juízo de que o par era o mais macio. Com base nisso, Carruthers adota uma visão de paridade eu/outro (ver Seção 2.1, acima) de nosso autoconhecimento de nossas atitudes, sustentando que só podemos introspectar, no sentido estrito, experiências conscientes como aquelas que surgem na percepção e na representação [*imagery*].

4.2.3 Da experiência consciente

A experiência consciente atualmente em curso - ou talvez imediatamente passada (se sustentarmos que o juízo introspectivo deve seguir temporalmente o estado ou processo introspectado, ou se enfatizarmos as preocupações levantadas na Seção 3.2 sobre o autoenfraquecimento do processo introspectivo) - é ao mesmo tempo o alvo mais universalmente reconhecido do processo introspectivo e o alvo mais comumente considerado ser conhecido com um alto grau de privilégio. Reivindicações de infalibilidade, indubitabilidade, incorrigibilidade e autorrevelação (ver Seção 4.1.1, acima) são mais comumente feitas para o autoconhecimento de estados como estar com dor ou ter uma experiência visual da cor vermelha, onde esses estados são interpretados como estados qualitativos, experiências subjetivas ou aspectos de nossa fenomenologia ou consciência. (Todos esses termos pretendem se referir indistintamente ao que Block [1995], Chalmers [1996] e outros filósofos contemporâneos chamam de “consciência fenomenal”). Se as atitudes são às vezes conscientes, então nós também poderíamos ser capazes de introspectar aquelas atitudes como parte da nossa capacidade de introspectar a experiência consciente em geral (Goldman 2006; Hill 2009).

É difícil estudar a precisão das autoatribuições de experiência consciente pelas mesmas razões que é difícil estudar a precisão das nossas autoatribuições de atitudes (Seção 4.2.2): não há medida amplamente aceita para superar ou confirmar o autorrelato. Na literatura médica sobre dor, por exemplo, nenhuma medida comportamental ou fisiológica da dor é geralmente considerada capaz de prevalecer sobre o autorrelato da dor atual, apesar do fato de que questões de escala continuam sendo um problema dentre e, especialmente, entre os indivíduos (Williams, Davies e Chadury 2000), assim como o é a avaliação retrospectiva (Redelmeier e Kahneman 1996). Quando os marcadores fisiológicos de dor e autorrelato se dissociam, não fica de modo algum claro que o marcador fisiológico deva ser considerado o índice mais preciso (para recomendações metodológicas, ver Price e Aydede 2005). Observações correspondentes se aplicam ao caso do prazer (Haybron 2008).

Conforme mencionado na Seção 3.3, acima, os primeiros psicólogos introspectivos afirmaram a dificuldade de introspectar com precisão a experiência consciente e alcançaram apenas um sucesso misto em suas tentativas de obter dados cientificamente replicáveis (e, portanto, presumivelmente precisos) por meio de uso de introspectadores treinados. Em alguns domínios eles alcançaram sucesso e replicabilidade consideráveis, como na

construção da “cor sólida” (uma representação das três dimensões primárias de variação na experiência de cor: matiz, saturação e luminosidade ou brilho), o mapeamento do tamanho de “diferenças apenas perceptíveis” entre as sensações e o limiar abaixo do qual um estímulo é muito fraco para ser experimentado, e a relação logarítmica (pelo menos aproximadamente) entre a intensidade de um estímulo sensorial e a intensidade da experiência resultante (a “lei de Weber-Fechner”). A psicofísica contemporânea - o estudo da relação entre estímulos físicos e as experiências sensoriais ou percepções resultantes - está enraizada nesses primeiros estudos introspectivos. No entanto, outros tipos de fenômenos se provaram resistentes ao consenso introspectivo entre laboratórios - tais como a possibilidade ou não de pensamento sem imagens (veja o verbete sobre “imagens mentais”¹²), a estrutura da emoção e os aspectos experienciais da atenção. Talvez esses fatos sobre o alcance do acordo introspectivo inicial e do desacordo aparentemente intratável lancem luz sobre o alcance em que a introspecção cuidadosa e bem treinada é e não é confiável.

¹² N. do E. Ver <<https://plato.stanford.edu/archives/win2019/entries/mental-imagery/>>

Ericsson e Simon (1984/1993; Ericsson 2003) discutem e revisam as relações entre o desempenho do participante em várias tarefas de resolução de problemas, suas verbalizações simultâneas de pensamentos conscientes (“protocolos de pensamento em voz alta”) e suas verbalizações imediatamente retrospectivas. A existência de boas relações nas direções previstas em muitas tarefas de resolução de problemas dá suporte empírico à visão de que os relatos das pessoas sobre seu fluxo de pensamentos geralmente refletem com precisão esses pensamentos. Por exemplo, Ericsson e Simon descobriram que relatórios retrospectivos e em voz alta de processos de pensamento se correlacionam com padrões previstos de movimento dos olhos e latência de respostas. Ericsson e Simon também citam estudos como o de Hamilton e Sanford (1978), que pediram aos participantes que fizessem juízos de sim ou não sobre se os pares de letras que estavam em ordem alfabética (como MO) ou não (como RP) e então descrevessem retrospectivamente seu método para chegar aos juízos. Quando os participantes relataram retrospectivamente saber a resposta “automaticamente” sem a intervenção de um processo consciente, os tempos de reação foram rápidos e não dependeram da distância entre as letras. Quando os participantes relataram retrospectivamente “percorrer” uma série sequencial de letras (como “LMNO”

quando incitado com “MO”), os tempos de reação se correlacionaram bem com a extensão relatada da passagem. Por outro lado, Flavell, Green e Flavell (1995) relatam erro introspectivo grosseiro e generalizado sobre o pensamento (consciente [*conscious*]) ocorrido recentemente e até mesmo atual em crianças pequenas; e Smallwood e Schooler (2006) revisam a literatura que sugere que as pessoas não são especialmente boas em detectar quando sua mente está divagando.

No século 20, os filósofos que argumentavam contra o infalibilismo frequentemente inventavam exemplos hipotéticos nos quais sugeriam que era plausível atribuir erro introspectivo; mas mesmo que tais exemplos sejam bem-sucedidos, eles são geralmente confinados a cenários improváveis, casos patológicos ou erros muito pequenos ou muito breves (por exemplo Armstrong 1963; Churchland 1988; Kornblith 1998, com referência à distinção entre erros sobre a experiência consciente atual e outros tipos de erros). No século 21, as críticas filosóficas da precisão dos juízos introspectivos sobre a consciência mudaram seu foco para casos de desacordo generalizado ou erro (putativo), seja entre pessoas comuns ou entre especialistas em pesquisa. Dannett (1991), Blackmore (2002) e Schwitzgebel (2011b), por exemplo, argumentam que a maioria das pessoas

está muito equivocada sobre a natureza da experiência da visão periférica. Esses autores argumentam que as pessoas experimentam clareza visual apenas em uma região pequena e em movimento rápido de cerca de 1 a 2 graus do arco visual, ao contrário da (dita) impressão generalizada que a maioria das pessoas têm de que experimentam uma faixa substancialmente mais ampla de clareza estável no campo visual. Outros argumentos recentes contra a precisão dos juízos introspectivos sobre a experiência consciente baseiam-se na citação do desacordo generalizado sobre se existe uma “fenomenologia do pensamento” além daquela da representação [*imagery*] e da emoção, sobre se a experiência sensorial como um todo é “rica” (incluindo, por exemplo, experiência tátil constante de seus pés em seus sapatos) ou “magro” (limitado principalmente apenas ao que está em atenção a qualquer momento), e sobre a natureza da experiência de representações [*imagery*] visuais (Hurlburt e Schwitzgebel 2007; Bayne e Spener 2010; Schwitzgebel 2011b; consultar Howwy 2011).

Irvine (2013, no prelo) argumentou que os problemas metodológicos nessa área são tão graves que o termo “consciência” [*consciousness*] deveria ser eliminado do discurso científico como impossível de ser efetivamente operacionalizado ou medido. Feest (2014) e Timmermans e Cleeremans (2015)

também destacam os desafios metodológicos substanciais do uso de relatos introspectivos na ciência da consciência, embora não sejam tão pessimistas quanto Irvine.

Referências Bibliográficas

- Alston, William P., 1971, “Varieties of privileged access”, *American Philosophical Quarterly*, 8: 223–241.
- Amedi, Amir, Rafael Malach, and Alvaro Pascual-Leone, 2005, “Negative BOLD differentiates visual imagery and perception”, *Neuron*, 48: 859–872.
- Aristotle, 3rd c. BCE/1961, *De Anima*, W.D. Ross (ed.), Oxford: Oxford University Press.
- Armstrong, David M., 1963, “Is introspective knowledge incorrigible?”, *Philosophical Review*, 72: 417–432.
- _____, 1968, *A materialist theory of the mind*, London: Routledge.
- _____, 1981, *The nature of mind and other essays*, Ithaca, NY: Cornell University Press.
- _____, 1999, *The mind-body problem*, Boulder, CO: Westview.
- Aronson, Elliot, 1968, “Dissonance theory: Progress and problems”, in *Theories of cognitive consistency*, Robert P. Abelson, et al. (eds.), Chicago: Rand McNally, 112–139.
- Aru, Jaan, Talis Bachmann, Wolf Singer, and Lucia Melloni, 2012, “Distilling the neural correlates of consciousness”, *Neuroscience and Biobehavioral Reviews*, 36: 737–746.

Augustinus, Aurelius, c. 420 C.E./1998, *The city of God against the pagans*, R.W. Dyson (tr.), Cambridge: Cambridge University Press.

Aydede, Murat, and Güven Güzeldere, 2005, “Cognitive architecture, concepts, and introspection: An information-theoretic solution to the problem of phenomenal consciousness”, *Noûs*, 39: 197–255.

Ayer, A.J., 1936 [1946], *Language, truth, and logic*, 2nd ed., London: Gollancz.

_____, 1963, *The concept of a person*, New York: St. Martin’s.

Baldwin, James Mark, 1901–1905, *Dictionary of philosophy and psychology*, New York: Macmillan.

Balog, Katalin, 2012, “Acquaintance and the mind-body problem”, in *New Perspectives on Type Identity*, Simone Gozzano & Christopher S. Hill (eds.), Cambridge: Cambridge University Press.

Bar-On, Dorit, 2004, *Speaking my mind*, Oxford: Oxford.

Barrett, Lisa Feldman, Batja Mesquita, Kevin N. Ochsner, and James J. Gross, 2007, “The experience of emotion”, *Annual Review of Psychology*, 58: 373–403.

Bayne, Tim, and Michelle Montague, eds., 2011, *Cognitive phenomenology*, Oxford: Oxford University Press.

Bayne, Tim, and Maja Spener, 2010, “Introspective humility”, *Philosophical Issues*, 20, 1–22.

- Bem, Daryl J., 1967, "Self-perception: An alternative interpretation of cognitive dissonance phenomena", *Psychological Review*, 74: 183-200.
- _____, 1972, "Self-perception theory", *Advances in Experimental Social Psychology*, 6: 1-62.
- Berkeley, George, 1710/1965, *A Treatise Concerning the Principles of Human Knowledge*, in *Principles, Dialogues, and Philosophical Correspondence*, Colin M. Turbayne (ed.), New York: Macmillan, 3-101.
- Bilgrami, Akeel, 2006, *Self-knowledge and resentment*, Cambridge, MA: Harvard University Press.
- Blackmore, Susan, 2002, "There is no stream of consciousness", *Journal of Consciousness Studies*, 9(5-6), 17-28.
- Block, Ned, 1995, "On a confusion about a function of consciousness", *Behavioral and Brain Sciences*, 18: 227-247.
- _____, 1996, "Mental paint and mental latex", *Philosophical Issues*, 7: 19-49.
- Boghossian, Paul, 1989, "Content and self-knowledge", *Philosophical Topics*, 17: 5-26.
- Bohner, Gerd, and Nina Dickel, 2011, "Attitudes and attitude change", *Annual Review of Psychology*, 62: 391-417.
- Borgoni, Cristina, 2016, "Dissonance and irrationality", *Pacific Philosophical Quarterly*, 97: 48-57.

Boring, Edwin G., 1921, "The stimulus-error", *American Journal of Psychology*, 32: 449-471.

_____, 1953, "A history of introspection", *Psychological Bulletin*, 50: 169-189.

Boyle, Matthew, 2009, "Two kinds of self-knowledge", *Philosophy & Phenomenological Research*, 78: 133-164.

_____, forthcoming, "Transparency and reflection", *Canadian Journal of Philosophy*.

Brascamp, Jan, Philipp Sterzer, Randolph Blake, and Tomas Knapen, 2017, "Multistable perception and the role of the frontoparietal cortex in perceptual inference", *Annual Review of Psychology*, 69: 77-103.

Brentano, Franz, 1874 [1995], *Psychology from an empirical standpoint*, 2nd English edition, Antos C. Rancurello, D. B. Terrell and Linda L. McAlister (trans.), New York: Routledge.

Burge, Tyler, 1979, "Individualism and the mental", *Midwest Studies in Philosophy*, 4: 73-121.

_____, 1988, "Individualism and self-knowledge", *Journal of Philosophy*, 85: 649-663.

_____, 1996, "Our entitlement to self-knowledge", *Proceedings of the Aristotelian Society*, 96: 91-116.

_____, 1998, "Reason and the first person", in *Knowing our own minds*, Crispin Wright, Barry C. Smith, and Cynthia Macdonald (eds.), Oxford: Oxford University Press, 243-270.

Byrne, Alex, 2018, *Transparency and self-knowledge*, Oxford: Oxford.

Campbell, Donald T., William H. Kruskal, and William P. Wallace, 1966, "Seating aggregation as an index of attitude", *Sociometry*, 29: 1-15.

Campbell, John, 1999, "Immunity to error through misidentification and the meaning of a referring term", *Philosophical Topics*, 25(1-2), 89-104.

Carruthers, Peter, 2005, *Consciousness: Essays from a higher-order perspective*, Oxford: Oxford University Press.

_____, 2011, *The opacity of mind*, Oxford: Oxford University Press.

Cassam, Quassim, 2014, *Self-knowledge for humans*, Oxford: Oxford University Press.

Cavell, Marcia, 2006, *Becoming a subject*, Oxford: Oxford University Press.

Chalmers, David J., 1996, *The conscious mind*, New York: Oxford.

_____, 2003, "The content and epistemology of phenomenal belief", in *Consciousness: New philosophical perspectives*, Quentin Smith and Aleksandar Jokic (eds.), Oxford: Oxford, 220-272.

Chapman, Dwight W., 1933, "Attensity, clearness, and attention", *American Journal of Psychology*, 45: 156-165.

Cheesman, Jim, and Philip M. Merikle, 1986, “Distinguishing conscious from unconscious perceptual processes”, *Canadian Journal of Psychology*, 40: 343-367.

Chisholm, Roderick M., 1981, *The first person*, Brighton, UK: Harvester.

Churchland, Paul M., 1988, *Matter and consciousness*, rev. ed., Cambridge, MA: MIT Press.

Coliva, Annalisa, 2016, *Varieties of self-knowledge*, London: Palgrave.

Comte, Auguste, 1830, *Cours de philosophie positive, volume 1*, Paris: Bacheleier, Libraire pour les Mathématiques.

Cooper, Joel, and Russell H. Fazio, 1984, “A new look at dissonance theory”, *Advances in Experimental Social Psychology*, 17: 229-266.

Csikszentmihalyi, Mihaly, 2014, *Flow and the Foundations of Positive Psychology*, Dordrecht: Springer.

Cui, Xu, Cameron B. Jeter, Dongni Yang, P. Read Montague, and David M. Eagleman, 2007, “Vividness of mental imagery: Individual variability can be measured objectively”, *Vision Research*, 47: 474-478.

Cunningham, William A., *et al.*, 2004, “Separable neural components in the processing of Black and White faces”, *Psychological Science*, 15: 806-813.

de Graaf, Tom A., Maartje c. de Jong, Rainer Goebel, Raymond van Ee, and Alexander T. Sack, 2011, “On the functional

relevance of frontal cortex for passive and volunatarily controlled bistable vision”, *Cerebral Cortex*, 21: 2322–2331.

de Graaf, Tom A., Po-Jang Hsieh, and Alexander T. Sack, 2012, “The ‘correlates’ in neural correlates of consciousness”, *Neuroscience and Biobehavioral Reviews*, 36: 191–197.

De Vaus, David, 1985/2002, *Surveys in social research*, London: Routledge.

Dehaene, Stanislaus, *et al.*, 2001, “Cerebral mechanisms of word masking and unconscious repetition priming”, *Nature Neuroscience*, 4: 752–758.

Dehaene, Stanislaus, and Jean-Pierre Changeux, 2011, “Experimental and theoretical approaches to conscious processing”, *Neuron*, 70: 200–227.

Del Cul, Antoine, Sylvain Baillet, and Stanislas Dehaene, 2007, “Brain dynamics underlying the nonlinear threshold for access to consciousness”, *PLoS Biology*, 5(10): e260).

Dennett, Daniel C., 1987, *The intentional stance*, Cambridge, MA: MIT Press.

_____, 1991, *Consciousness explained*, Boston: Little, Brown, and Co.

_____, 2000, “The case for rorts”, in *Rorty and his critics*, R.B. Brandom (ed.), Malden, MA: Blackwell, 91–101.

_____, 2002, “How could I be wrong? How wrong could I be?”, *Journal of Consciousness Studies*, 9(5–6): 13–6.

Descartes, René, 1637/1985, Discourse on the method, in *The philosophical writings of Descartes*, vol. 1, John Cottingham, Robert Stoothoff, and Dugald Murdoch (editors and translators), Cambridge: Cambridge University Press, 111–151.

_____, 1641/1984, Meditations on first philosophy, in *The philosophical writings of Descartes*, vol. 2, John Cottingham, Robert Stoothoff, and Dugald Murdoch (editors and translators), Cambridge: Cambridge University Press, 1–62.

Dovidio, John F., Kerry Kawakami, Craig Johnson, Brenda Johnson, and Adaiiah Howard, 1997, “On the nature of prejudice: Automatic and controlled processes”, *Journal of Experimental Social Psychology*, 33: 510–540.

Dretske, Fred, 1995, *Naturalizing the mind*, Cambridge, MA: MIT.

_____, 2004, “Knowing what you think vs. knowing that you think it”, in *The externalist challenge*, Richard Schantz (ed.), Berlin: Walter de Gruyter, 389–399.

Ebbinghaus, Hermann, 1885/1913, *Memory: A contribution to experimental psychology*, Henry A. Ruger and Clara E. Bussenius (translators), New York: Columbia.

Elster, Jon, 1983/2016, *Sour grapes*, Cambridge: Cambridge University Press.

Ericsson, K. Anders, 2003, “Valid and non-reactive verbalization of thoughts during performance of tasks: Towards a solution to the central problems of introspection as a source of scientific data”, *Journal of Consciousness Studies*, 10(9–10): 1–18.

Ericsson, K. Anders, and Herbert A. Simon, 1984/1993, *Protocol analysis*, revised edition, Cambridge, MA: MIT.

Evans, Gareth, 1982, *The varieties of reference*, John McDowell (ed.), Oxford: Clarendon; New York: Oxford University Press.

Fazio, Russell H., Joni R. Jackson, Bridget C. Dunton, and Carol J. Williams, 1995, "Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline?", *Journal of Personality and Social Psychology*, 69(6): 1013-1027.

Fechner, Gustav, 1860 [1964], *Elements of psychophysics*, Helmut E. Adler, Davis H. Howes, and Edwin G. Boring (ed. and trans.), New York: Holt, Rinehart, and Winston.

Feest, Uljana, 2014, "Phenomenal experiences, first-person methods, and the artificiality of experimental data", *Philosophy of Science*, 81: 927-939.

Fernández, Jorgi, 2003, "Privileged access naturalized", *Philosophical Quarterly*, 53: 352-372.

Festinger, Leon, 1957, *A theory of cognitive dissonance*, Stanford, CA: Stanford.

Festinger, Leon, and James M. Carlsmith, 1959, "Cognitive consequences of forced compliance", *Journal of Abnormal and Social Psychology*, 58: 203-210.

Flavell, John H., Frances L. Green, and Eleanor R. Flavell, 1995, *Young children's knowledge about thinking*, *Monographs of the Society for Research in Child Development*, 60(1).

Fodor, Jerry A., 1983, *Modularity of mind*, Cambridge, MA: MIT.

_____, 1998, *Concepts: Where cognitive science went wrong*, Oxford: Oxford University Press.

Frässle, Stefan, Jens Sommer, Andreas Jansen, Marnix Naber, and Wolfgang Einhäuser, 2014, “Binocular rivalry: Frontal activity relates to introspection and action but not to perception”, *Journal of Neuroscience*, 34: 1738–1747

Funder, David C., 1999, *Personality judgment*, London: Academic.

Gallois, Andre, 1996, *The world without, the mind within*, Cambridge: Cambridge.

Galton, Francis, 1869/1891, *Hereditary genius*, rev. ed., New York: Appleton.

Gardner, Sebastian, 1993, *Irrationality and the philosophy of psychoanalysis*, Cambridge: Cambridge University Press.

Gazzaniga, Michael S., 1995, “Consciousness and the cerebral hemispheres”, in *The Cognitive Neurosciences*, Michael S. Gazzaniga (ed.), Cambridge, MA: MIT, 1391–1400.

Gawronski, Bertram, and Galen V. Bodenhausen, 2006, “Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change”, *Psychological Bulletin*, 132: 692–731.

Gawronski, Bertram, Mike Morrison, Curtis E. Phillips, and Silvia Galdi, 2017, “Temporal stability of implicit and explicit

measures: A longitudinal analysis”, *Personality and Social Psychology Bulletin*, 43: 300–312.

Gendler, Tamar Szabó, 2008a, “Alief and belief”, *Journal of Philosophy*, 105: 634–663.

_____, 2008b, “Alief in action, and reaction”, *Mind & Language*, 23: 552–585.

Gennaro, Rocco J., 1996, *Consciousness and Self-Consciousness*, Amsterdam: John Benjamins.

Gertler, Brie, 2000, “The mechanics of self-knowledge”, *Philosophical Topics*, 28: 125–146.

_____, 2001, “Introspecting phenomenal states”, *Philosophy and Phenomenological Research*, 63: 305–328.

_____, 2011, “Self-knowledge and the transparency of belief”, in *Self-knowledge*, Anthony Hatzimoysis (ed.), Oxford: Oxford University Press.

Goldman, Alvin I., 1989, “Interpretation psychologized”, *Mind and Language*, 4: 161–185.

_____, 2000, “Can science know when you’re conscious?”, *Journal of Consciousness Studies*, 7(5): 3–22.

_____, 2006, *Simulating minds*, Oxford: Oxford.

Gopnik, Alison, 1993a, “How we know our minds: The illusion of first-person knowledge of intentionality”, *Behavioral and Brain Sciences*, 16: 1–14.

_____, 1993b, “Psychopsychology”, *Consciousness and Cognition*, 2: 264–280.

Gopnik, Alison, and Andrew N. Meltzoff, 1994, “Minds, bodies and persons: Young children’s understanding of the self and others as reflected in imitation and ‘theory of mind’ research”, in *Self-awareness in animals and humans*, Sue Taylor Parker, Robert W. Mitchell, and Maria L. Boccia (eds.), New York: Cambridge, 166–186.

Gordon, Robert M., 1995, “Simulation without introspection or inference from me to you”, in *Mental simulation*, Martin Davies and Tony Stone (eds.), Oxford: Blackwell.

_____, 2007, “Ascent routines for propositional attitudes”, *Synthese*, 159: 151–165.

Green, David M., and John A. Swets, 1966, *Signal detection theory and psychophysics*, Oxford: Wiley.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L.K. Schwartz, 1998, “Measuring individual differences in implicit cognition: The Implicit Association Test”, *Journal of Personality and Social Psychology*, 74: 1464–1480.

Greenwald, Anthony G., and Brian A. Nosek, 2009, “Attitudinal dissociation: What does it mean?”, in *Attitudes: Insights from the New Implicit Measures*, Richard E. Petty, Russell H. Fazio, and Pablo Briñol (eds.), New York: Taylor and Francis, 65–82.

Hahn, Adam, Charles M. Judd, Holen K. Hirsh, and Irene V. Blair, 2014, “Awareness of implicit attitudes”, *Journal of Experimental Psychology: General*, 143: 1369–1392.

Hamilton, Andy, 2007, “Memory and self-consciousness: Immunity to error through misidentification”, *Synthese*, 171: 409–417.

Hamilton, J.M.E., and A.J. Sanford, 1978, “The symbolic distance effect for alphabetic order judgements: A subjective report and reaction time analysis”, *Quarterly Journal of Experimental Psychology*, 30: 33–43.

Harman, Gilbert, 1990, “The intrinsic quality of experience”, in James Tomberlin, (ed.), *Philosophical Perspectives*, 4, Atascadero, CA: Ridgeview, 31–52.

Hart, Allen J., Paul J. Whalen, Lisa M. Shin, Sean C. McInerney, Hakan Fischer, and Scott L. Rauch, 2000, “Differential response in the human amygdala to racial outgroup vs ingroup face stimuli”, *NeuroReport*, 11: 2351–2355.

Haybron, Daniel M., 2008, *The pursuit of unhappiness*, Oxford: Oxford University Press.

Heal, Jane, 2002, “On first-person authority”, *Proceedings of the Aristotelian Society*, 102: 1–19.

Heil, John, 1988, “Privileged access”, *Mind*, 97: 238–251.

Helmholtz, Hermann, 1856/1962, *Helmholtz’s Treatise on Physiological Optics*, James P.C. Southall (ed.), New York: Dover. [Translation based on 1924 edition.]

Hill, Christopher S., 1991, *Sensations: A defense of type materialism*, Cambridge: Cambridge University Press.

_____, 2009, *Consciousness*, Cambridge: Cambridge University Press.

Hirstein, William, 2005, *Brain fiction*, Cambridge, MA: MIT.

Hohwy, Jakob, 2011, “Phenomenal variability and introspective reliability”, *Mind & Language*, 26: 261–286.

Horgan, Terence, John L. Tienson, and George Graham, 2006, “Internal-world skepticism and mental self-presentation”, in *Self-representational approaches to consciousness*, Uriah Kriegel and Kenneth Williford (eds.), Cambridge, MA: MIT, 191–207.

Horgan, Terence, and Uriah Kriegel, 2007, “Phenomenal epistemology: What is consciousness that we may know it so well?”, *Philosophical Issues*, 17(1): 123–144.

Hume, David, 1739 [1978], *A treatise of human nature*, L.A. Selby-Bigge and P.H. Nidditch (eds.), Oxford: Clarendon.

_____, 1748/1975, An enquiry concerning human understanding, in *Enquiries concerning human understanding and concerning the principles of morals*, L.A. Selby-Bigge and P.H. Nidditch (eds.), Oxford: Clarendon, 1–165.

Humphrey, George, 1951, *Thinking: An introduction to its experimental psychology*, London: Methuen.

Hurlburt, Russell T., 1990, *Sampling normal and schizophrenic inner experience*, New York: Plenum.

Hurlburt, Russell T., 2011, *Investigating pristine inner experience*, Cambridge: Cambridge.

Hurlburt, Russell T., and Christopher L. Heavey, 2006, *Exploring inner experience*, Amsterdam: John Benjamins.

Hurlburt, Russell T., and Eric Schwitzgebel, 2007, *Describing inner experience? Proponent meets skeptic*, Cambridge, MA: MIT.

Husserl, Edmund, 1913 [1982], *Ideas*, Book I, T.E. Klein and W.E. Pohl (trs.), Dordrecht: Kluwer.

Ilg, Rüdiger, Afra M. Wohlschläger, Stefan Burazanis, Andreas Wöller, Sabine Nunnemann, and Mark Mühlau, 2008, “Neural correlates of spontaneous percept switches in ambiguous stimuli: An event-related functional magnetic resonance imaging study”, *European Journal of Neuroscience*, 28: 2325–2332.

Irvine, Elizabeth, 2013, *Consciousness as a scientific concept*, Dordrecht: Springer.

_____, forthcoming, “Developing dark pessimism towards the justificatory role of introspective reports”, *Erkenntnis*.

Ito, Tiffany A., and John T. Cacioppo, 2007, “Attitudes as mental and neural states of readiness”, in *Implicit measures of attitudes*, Bernd Wittenbrink and Norbert Schwarz (eds.), New York: Guilford, 125–158.

Jack, Anthony, and Andreas Roepstorff, 2003, Trusting the subject, vol. 1, special issue of the *Journal of Consciousness Studies*, no. 10(9–10).

_____, 2004, Trusting the subject, vol. 2, special issue of the *Journal of Consciousness Studies*, 11(7–8).

James, William, 1890 [1981], *The principles of psychology*, Cambridge, MA: Harvard.

Jaynes, Julian, 1976, *The origin of consciousness in the breakdown of the bicameral mind*, New York: Houghton Mifflin.

Johansson, Petter, Lars Hall, Sverker Sikström, and Andreas Olsson, 2005, "Failure to detect mismatches between intention and outcome in a simple decision task", *Science*, 310: 116-119.

Johansson, Petter, Lars Hall, Sverker Sikström, Betty Tärning, and Andreas Lind, 2006, "How something can be said about telling more than we can know: On choice blindness and introspection", *Consciousness and Cognition*, 15: 673-692.

Kamphuisen, Allard, Markus Bauer, and Raymond van Ee, 2008, "No evidence for widespread synchronized networks in binocular rivalry: MEG frequency tagging entrains primary early visual cortex", *Journal of Vision*, 8(5): article 4.

Kant, Immanuel, 1781/1997, *The critique of pure reason*, Paul Guyer and Allen W. Wood (eds. and trs.), Cambridge: Cambridge.

Kay, Aaron C., Maria C. Jimenez, and John T. Jost, 2002, "Sour grapes, sweet lemons, and the anticipatory rationalization of the status quo", *Personality and Social Psychology Bulletin*, 28: 1300-1312.

Kihlstrom, John F., "Implicit methods in social psychology", in *The SAGE handbook of methods in social psychology*, Carol Sansone, Carolyn C. Morf, and A.T. Panter (eds.), Thousand Oaks, CA: Sage, 195-212.

Kind, Amy, 2003, “What’s so transparent about transparency?”, *Philosophical Studies*, 115: 225-244.

Kleinschmidt, A., C. Büchel, S. Zeki, and R.S.J. Frackowiak, 1998, “Human brain activity during spontaneously reversing perception of ambiguous figures”, *Proceedings of the Royal Society B*, 265: 2427-2433.

Knapen, Tomas, Jan Brascamp, Joel Pearson, Raymond van Ee, and Randolph Blake, 2011, “The role of frontal and parietal areas in bistable perception”, *Journal of Neuroscience*, 31: 10293-10301.

Koch, Christof, Marcello Massimini, Melanie Boly, and Giulio Tononi, 2016, “Neural correlates of consciousness: progress and problems”, *Nature Reviews Neuroscience*, 17: 307-321.

Kornblith, Hilary, 1998, “What is it like to be me?”, *Australasian Journal of Philosophy*, 76: 48-60.

Kosslyn, Stephen M., Daniel Reisberg, and Marlene Behrmann, 2006, “Introspection and mechanism in mental imagery”, in *The Dalai Lama at MIT*, Anne Harrington and Arthur Zajonc (eds.), Cambridge, MA: Harvard, 79-90.

Kriegel, Uriah, 2009, *Subjective consciousness*, Oxford: Oxford.

Külpe, Oswald, 1893 [1895], *Outlines of psychology*, Edward Titchener (trans.), London: George Allen & Unwin.

Kusch, Martin, 1999, *Psychological knowledge*, London, Routledge.

Lambie, John A., and Anthony J. Marcel, 2002, "Consciousness and the varieties of emotion experience: A theoretical framework", *Psychological Review*, 109: 219-259.

Lane, Kristin A., Mahzarin R. Banaji, Brian A. Nosek, and Anthony G. Greenwald, 2007, "Understanding and using the Implicit Association Test: IV", in *Implicit measures of attitudes*, Bernd Wittenbrink and Norbert Schwarz (eds.), New York: Guilford, 59-102.

Larson, Reed, and Mihaly Csikszentmihalyi, 1983, "The Experience Sampling Method" in Harry T. Reis, (ed.), *Naturalistic approaches to studying social interaction*, San Francisco: Jossey-Bass, 41-56.

Lear, Jonathan, 1998, *Open-minded*, Cambridge, MA: Harvard.

Levy, Neil, 2015, "Neither fish nor fowl: Implicit attitudes as patchy endorsements", *Noûs*, 49: 800-823.

Lewis, C.I., 1946, *An analysis of knowledge and valuation*, La Salle, IL: Open Court.

Locke, John, 1690 [1975], *An essay concerning human understanding*, Peter H. Nidditch (ed.), Oxford: Oxford University Press.

Lumer, Erik D., Karl J. Friston, and Geraint Rees, 1998, "Neural correlates of perceptual rivalry in the human brain", *Science*, 280: 1930-1934.

Lycan, William G., 1996, *Consciousness and experience*, Cambridge, MA: MIT.

Lyons, William, 1986, *The disappearance of introspection*, Cambridge, MA: MIT.

Lyubomirsky, Sonja, and Lee Ross, 1999, “Changes in attractiveness of elected, rejected, and precluded alternatives: A comparison of happy and unhappy individuals”, *Journal of Personality and Social Psychology*, 76: 988–1007.

Machery, Edouard, 2016, “De-Freuding implicit attitudes”, in *Implicit Bias and philosophy*, volume 1: Metaphysics and epistemology, Michael Brownstein and Jennifer Saul (eds.), Oxford: Oxford University Press.

Macmillan, Neil A., and C. Douglas Creelman, 1991, *Detection theory*, Cambridge: Cambridge University Press.

Mandelbaum, Eric, 2016. “Attitude, inference, association: On the propositional structure of implicit bias”, *Noûs*, 50: 629–658.

Marks, David F., 1985, “Imagery paradigms and methodology” *Journal of Mental Imagery*, 9: 93–105.

Marr, David, 1983, *Vision*, New York: Freeman.

Martin, Michael G.F., 2002, “The transparency of experience”, *Mind and Language*, 17: 376–425.

Maudsley, Henry, 1867 [1977], *Physiology and pathology of the mind*, Daniel N. Robinson (ed.), Washington, DC: University Publications of America.

McGeer, Victoria, 1996, “Is ‘self-knowledge’ an empirical problem? Renegotiating the space of philosophical explanation”, *Journal of Philosophy*, 93: 483–515.

_____, 2008, “The moral development of first-person authority”, *European Journal of Philosophy*, 16: 81-108.

McGeer, Victoria, and Philip Pettit, 2002, “The self-regulating mind”, *Language and Communication*, 22: 281-299.

Megumi, Fukuda, Bahador Bahrami Ryota Kanai, and Geraint Rees, 2015, “Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception”, *NeuroImage*, 107: 190-197.

Mele, Alfred, 2001, *Self-deception unmasked*, Princeton, NJ: Princeton.

Mengzi, 3rd c. BCE [2008], TITLE, B.W. Van Norden (tr.), Indianapolis: Hackett.

Merickle, Philip M., Daniel Smilek, and John D. Eastwood, 2001, “Perception without awareness: Perspectives from cognitive psychology”, *Cognition*, 79: 115-134.

Mill, James, 1829 [1878], *Analysis of the Phenomena of the Human Mind*, John Stuart Mill (ed.), London: Longmans, Green, Reader, and Dyer.

Mill, John Stuart, 1865 [1961], *Auguste Comte and positivism*, Ann Arbor, MI: University of Michigan.

Mole, Christopher, 2011, *Attention is cognitive unison*, Oxford: Oxford University Press.

Moll, Albert, 1889 [1911], *Hypnotism*, Arthur F. Hopkirk (ed.), New York: Charles Scribner’s Sons.

Moore, George Edward, 1903, "The refutation of idealism", *Mind*, 12: 433-453.

_____, 1942, "A reply to my critics", in *The philosophy of G.E. Moore*, in P.A. Schilpp (ed.), New York: Tudor, 535-677.

_____, 1944 [1993], "Moore's paradox", in *G.E. Moore, Selected writings*, Thomas Baldwin (ed.), London: Routledge, 207-212.

Moran, Richard, 2001, *Authority and estrangement*, Princeton: Princeton.

Müller, G.E., 1904, *Die Gesichtspunkte und die Tatsachen der psychophysischen Methodik*, Wiesbaden: J.F. Bergmann.

Nahmias, Eddy, 2002, "Verbal reports on the contents of consciousness: Reconsidering introspectionist methodology", *Psyche*, 8(21).

Nichols, Shaun, and Stephen P. Stich, 2003, *Mindreading*, Oxford: Oxford University Press.

Nisbett, Richard E., and Nancy Bellows, 1977, "Verbal reports about causal influences on social judgments: Private access versus public theories", *Journal of Personality and Social Psychology*, 35: 613-624.

Nisbett, Richard E., and Lee Ross, 1980, *Human inference*, Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, Richard E., and Timothy DeCamp Wilson, 1977, "Telling more than we can know: Verbal reports on mental processes", *Psychological Review*, 84: 231-259.

Noë, Alva, 2004, *Action in perception*, Cambridge, MA: MIT Press.

Noë, Alva, and Evan Thompson, 2004, “Are there neural correlates of consciousness?”, *Journal of Consciousness Studies*, 11: (1): 3–28.

Oswald, Frederick, Gregory Mitchell, Hart Blanton, James Jaccard, and Philip Tetlock, 2013, “Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies”, *Journal of Personality and Social Psychology*, 105: 171–192.

Overgaard, Morten, Kristian Sandberg, and Mads Jensen, 2008, “The neural correlate of consciousness?”, *Journal of Theoretical Biology*, 254: 713–715.

Papineau, David, 2002, *Thinking about consciousness*, Oxford: Oxford University Press.

Parkkonen, Lauri, Jesper Andersson, Matti Hämäläinen, and Riitta Hari, 2008, “Early visual brain areas reflect the percept of an ambiguous scene”, *Proceedings of the National Academy of Sciences*, 105: 20500–20504.

Paul, Elliot S., 2018, “Descartes's anti-transparency and the need for radical doubt”, *Ergo*, 5: 1083–1129.

Paulhus, Delroy L., and Oliver P. John, 1998, “Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives”, *Journal of Personality*, 66: 1025–1060.

Payne, B. Keith, Heidi A. Vuletich and Kristjen B. Lundberg, 2017, "The bias of crowds: How implicit bias bridges personal and systemic prejudice", *Psychological Inquiry*, 28: 233-248.

Peacocke, Christopher, 1998, "Conscious attitudes, attention, and self-knowledge", in *Knowing our own minds*, Crispin Wright, Barry C. Smith, and Cynthia Macdonald (eds.), Oxford: Oxford University Press, 63-99.

Petitmengin, Claire, 2006, "Describing one's subjective experience in the second person: An interview method for the science of consciousness", *Phenomenology and the Cognitive Sciences*, 5: 229-269.

Petty, Richard E., Russell H. Fazio, and Pablo Briñol (eds.), 2009, *Attitudes: Insights from the new implicit measures*, New York: Taylor and Francis.

Phillips Ian, 2018, "The methodological puzzle of phenomenal consciousness", *Philosophical Transactions of the Royal Society B*, 373: 20170347.

Pillsbury, W.B., 1908, *Attention*, London: Swan Sonnenschein.

Price, Donald D., and Murat Aydede, 2005, "The experimental use of introspection in the scientific study of pain and its integration with third-person methodologies: The experiential-phenomenological approach", in *Pain: New essays on its nature and the methodology of its study*, Murat Aydede (ed.), Cambridge, MA: MIT, 243-273.

Prinz, Jesse, 2004, "The fractionation of introspection", *Journal of Consciousness Studies*, 11(7-8): 40-57.

_____, 2007, “Mental pointing: Phenomenal knowledge without concepts”, *Journal of Consciousness Studies*, 14(9-10): 184-211.

_____, 2012, *The conscious brain*, Oxford: Oxford.

Pryor, James, 1999, “Immunity to error through misidentification”, *Philosophical Topics*, 26(1-2): 271-304.

Putnam, Hilary, 1975, “The meaning of ‘meaning’” in *Hilary Putnam, Philosophical papers*, vol. 2, Cambridge: Cambridge University Press, 215-271.

Quiroga, R. Quian, R. Mukamel, E.A. Isham, and I. Fried, 2008, “Human single-neuron responses at the threshold of conscious recognition”, *Proceedings of the National Academy of Sciences*, 105: 3599-3604.

Redelmeier, Donald A., and Daniel Kahneman, 1996, “Patients’ memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures”, *Pain*, 66: 3-8.

Rees, Geraint, and Chris Frith, 2007, “Methodologies for identifying the neural correlates of consciousness”, in *The Blackwell Companion to Consciousness*, Max Velmans and Susan Schneider (eds.), Malden, MA: Blackwell, 553-566.

Richet, Charles, 1884, *L’homme et l’intelligence*, Paris: F. Alcan.

Rodriguez, Eugenio, Nathalie George, Jean-Philippe Lachauz, Jacques Martinerie, Bernard Renault, and Francisco J. Varela,

1999, “Perception’s shadow: Long-distance synchronization of human brain activity”, *Nature*, 397: 430–433.

Rorty, Richard, 1970, “Incorrigibility as the mark of the mental”, *Journal of Philosophy*, 67: 399–424.

Rosenthal, David M., 1990, “Two concepts of consciousness”, *Philosophical Studies*, 49: 329–359

_____, 2001, “Introspection and self-interpretation”, *Philosophical Topics*, 28(2): 201–233.

_____, 2005, *Consciousness and Mind*, Oxford: Oxford University Press.

Ryle, Gilbert, 1949, *The concept of mind*, New York: Barnes and Noble.

Samoilova, Kateryna, 2016, “Transparency and introspective unification”, *Synthese*, 193: 3363–3381.

Salti, Moti, *et al.*, 2015, “Distinct cortical codes and temporal dynamics for conscious and unconscious percepts”, *eLife*, 4: e05652.

Sandberg, Kristian, Bahador Bahrami, Ryota Kanai, Gareth Robert Barnes, Morten Overgaard, and Geraint Rees, 2013, “Early visual responses predict conscious face perception within and between subjects during binocular rivalry”, *Journal of Cognitive Neuroscience*, 25: 969–985.

Schwitzgebel, Eric, 2002, “A phenomenal, dispositional account of belief”, *Noûs*, 36: 249–275.

_____, 2005, “Difference tone training”, *Psyche*, 11(6).

_____, 2007, “No unchallengeable epistemic authority, of any sort, regarding our own conscious experience-*contra* Dennett?”, *Phenomenology and the Cognitive Sciences*, 6: 107-113.

_____, 2010, “Acting contrary to our professed beliefs, or the gulf between occurrent judgment and dispositional belief”, *Pacific Philosophical Quarterly*, 91: 531-553.

_____, 2011a, “Knowing your own beliefs”, *Canadian Journal of Philosophy*, 35 (Supplement: Belief and Agency, ed. D. Hunter): 41-62.

_____, 2011b, *Perplexities of consciousness*, Cambridge, MA: MIT.

_____, 2012, “Introspection, what?”, in *Introspection and consciousness*, Declan Smithies and Daniel Stoljar (eds.), Oxford: Oxford.

_____, forthcoming, “The pragmatic metaphysics of belief”, in *The fragmented mind*, Cristina Borgoni and Dirk Kindermann (eds.), Oxford: Oxford University Press.

Scollon, Christie Napa, Ed Diener, Shigehiro Oishi, Robert Biswas-Diener, 2005, “An experience-sampling and cross-cultural investigation of the relation between pleasant and unpleasant affect”, *Cognition and Emotion*, 19: 27-52.

Searle, John R., 1983, *Intentionality*, Cambridge: Cambridge.

_____, 1992, *The rediscovery of the mind*, Cambridge, MA: MIT Press.

Shoemaker, Sydney, 1963, *Self-knowledge and self-identity*, Ithaca, NY: Cornell University Press.

_____, 1968, “Self-reference and self-awareness”, *Journal of Philosophy*, 65: 555–567.

_____, 1988, “On knowing one’s own mind”, *Philosophical Perspectives*, 2: 183–209.

_____, 1994a, “Self-knowledge and ‘inner sense’. Lecture I: The object perception model”, *Philosophy and Phenomenological Research*, 54: 249–269.

_____, 1994b, “Self-knowledge and ‘inner sense’. Lecture II: The broad perceptual model”, *Philosophy and Phenomenological Research*, 54: 271–290.

_____, 1994c, “Self-knowledge and ‘inner sense’. Lecture III: The phenomenal character of experience”, *Philosophy and Phenomenological Research*, 54: 291–314.

_____, 1995, “Moore’s paradox and self-knowledge”, *Philosophical Studies*, 77: 211–228.

_____, 2012, “Self-intimation and second-order belief”, in *Introspection and consciousness*, Declan Smithies and Daniel Stoljar (eds.), Oxford: Oxford.

Siewert, Charles, 2004, “Is experience transparent?”, *Philosophical Studies*, 117: 15–41.

_____, 2012, “On the phenomenology of introspection”, in *Introspection and consciousness*, Declan Smithies and Daniel Stoljar (eds.), Oxford: Oxford.

Singal, Jesse, 2017, “Psychology’s favorite tool for measuring racism isn’t up to the job”, *The Cut*, available online.

Sirken, Monroe G., Douglas J. Herrmann, Susan Schechter, Norbert Schwarz, Judith N. Tanur, Roger Tourangeau (eds.), 1999, *Cognition and survey research*, New York: John Wiley and Sons.

Smallwood, Jonathan, and Jonathan W. Schooler, 2006, “The restless mind”, *Psychological Bulletin*, 132: 946-958.

Smith, A.D., 2008, “Translucent experiences”, *Philosophical Studies*, 140:197-212.

Spener, Maja, forthcoming, “Disagreement about cognitive phenomenology”, in *Cognitive phenomenology*, Tim Bayne and Michelle Montague (eds.), Oxford: Oxford University Press.

Stoljar, Daniel, 2004, “The argument from diaphanousness”, in *New essays in the philosophy of language and mind*, Maite Ezcurdia, Robert J. Stainton, and Christopher Viger (eds.), Calgary: University of Calgary, 341-390.

Stone, Jeff, and Joel Cooper, 2001, “A self-standards model of cognitive dissonance”, *Journal of Experimental Social Psychology*, 37: 228-243.

Summerfield, Christopher, Anthony Ian Jack, and Adrian Philip Burgess, 2002, “Induced gamma activity is associated with

conscious awareness of pattern masked nouns”, *International Journal of Psychophysiology*, 44: 93-100.

Taylor, Shelley E., and Jonathon D. Brown, 1988, “Illusion and well-being: A social psychological perspective on mental health”, *Psychological Bulletin*, 103: 193-210.

Thomas, Nigel, 1999, “Are theories of imagery theories of imagination?”, *Cognitive Science*, 23: 207-245.

Timmermans, Bert, and Alex Cleeremans, 2015, “How can we measure awareness? An overview of current methods”, in *Behavioural methods in consciousness research*, Morten Overgaard (ed.), Oxford: Oxford University Press.

Titchener, E.B., 1901-1905, *Experimental psychology*, New York: Macmillan.

_____, 1908 [1973], *Lectures on the elementary psychology of feeling and attention*, New York: Arno.

_____, 1912a, “Prolegomena to a study of introspection”, *American Journal of Psychology*, 23: 427-448.

_____, 1912b, “The schema of introspection”, *American Journal of Psychology*, 23: 485-508.

Tong, Frank, Ming Meng, and Randolph Blake, 2006, “Neural bases of binocular rivalry”, *Trends in Cognitive Sciences*, 10: 502-511.

Tong, Frank, Ken Nakayama, J. Thomas Vaughan, and Nancy Kanwisher, 1998, “Binocular rivalry and visual awareness in human extrastriate cortex”, *Neuron*, 21: 753-759.

Tononi, Giulio, and Christof Koch, 2008, “The neural correlates of consciousness: An update”, *Annals of the New York Academy of Sciences: The Year in Cognitive Neuroscience 2008*, 1124: 239–261.

Tononi, Giulio, Ramesh Srinivasan, D. Patrick Russell, and Gerald M. Edelman, 1998, “Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses”, *Proceedings of the National Academy of Sciences*, 95: 3198–3203.

Tsuchiya, Naotsugu, Melanie Wilke, Stefan Frässle, and Victor A.F. Lamme, 2015, “No-report paradigms: Extracting the true neural correlates of consciousness”, *Trends in Cognitive Sciences* 19: 757–770

Tye, Michael, 1995, *Ten problems about consciousness*, Cambridge, MA: MIT.

_____, 2000, *Consciousness, color, and content*, Cambridge, MA: MIT.

_____, 2002, “Representationalism and the transparency of experience”, *Noûs*, 36: 137–151.

_____, 2009, *Consciousness revisited*, Cambridge, MA: MIT.

Van Gulick, Robert, 1993, “Understanding the phenomenal mind: Are we all just armadillos?”, in *Consciousness: Psychological and philosophical essays*, Martin Davies and Glyn W. Humphreys (eds.), Oxford: Blackwell, 134–154.

Vanman, Eric J., Brenda Y. Paul, Tiffany A. Ito, and Norman Miller, 1997, "The modern face of prejudice and structural features that moderate the effect of cooperation on affect" *Journal of Personality and Social Psychology*, 73: 941-959.

Varela, Francisco J., 1996, "Neurophenomenology: A methodological remedy for the hard problem", *Journal of Consciousness Studies*, 3(4): 330-49.

Vazire, Simine, 2010, "Who knows what about a person? The Self-Other Knowledge Asymmetry (SOKA) model", *Journal of Personality and Social Psychology*, 98: 281-300.

Velleman, J. David, 2000, *The possibility of practical reason*, Oxford: Oxford University Press.

Watson, John B., 1913, "Psychology as the behaviorist views it", *Psychological Review*, 20: 158-177.

Wegner, Daniel M., 2002, *The illusion of conscious will*, Cambridge, MA: MIT.

Wegner, Daniel M. and Thalia Wheatley, 1999, "Apparent mental causation", *American Psychologist*, 54: 480-492.

White, Peter A., 1988, "Knowing more about what we can tell: 'Introspective access' and causal report accuracy ten years later", *British Journal of Psychology*, 79: 13-45.

Williams, Amanda C. de C., Huw Talfryn Oakley Davies, and Yasmin Chadury, 2000, "Simple pain rating scales hide complex idiosyncratic meanings", *Pain*, 85: 457-463.

Wilson, Timothy D., 2002, *Strangers to ourselves*, Cambridge, MA: Harvard.

Wilson, Timothy D., Samuel Lindsey, and Tonya T. Schooler, 2000, "A model of dual attitudes", *Psychological Review*, 107: 101-126.

Wittenbrink, Bernd, Charles M. Judd, and Bernadette Park, 1997, "Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures", *Journal of Personality and Social Psychology*, 72: 262-274.

Wittenbrink, Bernd, and Norbert Schwarz (eds.), 2007, *Implicit measures of attitudes*, New York: Guilford.

Wittgenstein, Ludwig, 1953/1968, *Philosophical investigations*, 3rd edition, G.E.M. Anscombe (translator), New York: Macmillan.

Wollheim, Richard, 1981, *Sigmund Freud*, New York: Cambridge.

_____, 2003, "On the Freudian unconscious", *Proceedings and Addresses of the American Philosophical Association*, 77(2): 23-35.

Wright, Crispin, 1989, "Wittgenstein's later philosophy of mind: Sensation, privacy, and intention", *Journal of Philosophy*, 86: 622-634.

_____, 1998, "Self-knowledge: The Wittgensteinian legacy", in *Knowing our own minds*, Crispin Wright, Barry C. Smith, and Cynthia Macdonald (eds.), Oxford: Oxford.

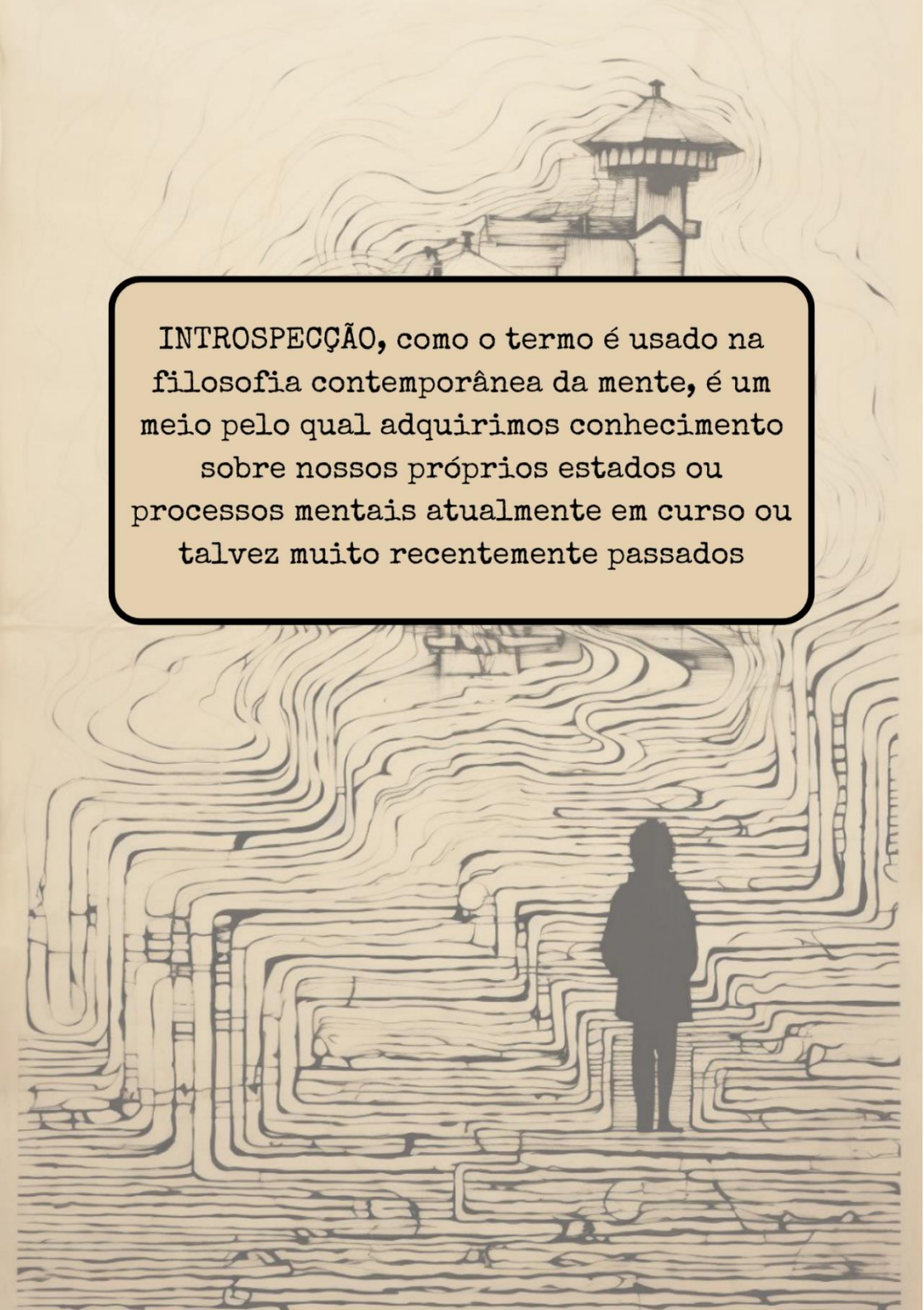
Wundt, Wilhelm, 1874 [1908], *Grundzüge der physiologischen Psychologie* (6th ed.), Leipzig: Wilhelm Engelmann.

_____, 1888, Selbstbeobachtung und innere Wahrnehmung, *Philosophische Studien*, 4: 292–309.

_____, 1896 [1902], *Outlines of psychology* (4th ed.), 2nd English ed., Charles Hubbard Judd (trans.), Leipzig: Wilhelm Engelmann.

_____, 1907, “Über Ausfrageexperimente und über die Methoden zur Psychologie des Denkens”, *Psychologische Studien*, 3: 301–360.

Zimmerman, Aaron, 2018, *Belief: A pragmatic picture*, Oxford: Oxford University Press.



INTROSPECÇÃO, como o termo é usado na filosofia contemporânea da mente, é um meio pelo qual adquirimos conhecimento sobre nossos próprios estados ou processos mentais atualmente em curso ou talvez muito recentemente passados