

# Uma análise do auto-engano fatal

Antonio Trajano Menezes Arruda

**Como citar:** ARRUDA, A. T. M. Uma análise do auto-engano fatal. *In:* GONZALES, M. E. Q.; BROENS, M. C. (org.). **Encontro com as Ciências Cognitivas**. Marília: Unesp Marília Publicações, 1998. 2 v. p. 167-186. DOI: <https://doi.org/10.36311/1998.85-86738-03-4.p167-186>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 (CC BY-NC-ND 4.0).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-No comercial-Sin derivados 4.0 (CC BY-NC-ND 4.0).

## UMA ANÁLISE DO AUTO-ENGANO FATUAL

Antonio Trajano Menezes ARRUDA<sup>1</sup>

### 1 Posição do problema

O auto-engano é um estado no qual a pessoa se põe a si mesma e não uma condição que a afeta à maneira de uma doença em relação à qual ela é totalmente passiva e não tem nenhuma responsabilidade. A definição do conceito geral de auto-engano utilizada na presente análise é a seguinte: um erro é um caso de auto-engano se e somente se ele é uma instância de **erro cognitivo motivado**, onde o motivo é específico do juízo em relação ao qual o sujeito se diz estar em auto-engano.

O erro cognitivo motivado é a coisa por excelência que nos ensina a respeito de uma importante dimensão particular da falibilidade da razão, a saber, aquela em relação à qual o intelecto da pessoa erradamente vem a enxergar o mundo de uma certa maneira essencialmente porque essa pessoa quer que o mundo seja dessa maneira.

Há, sobretudo em Filosofia mas também em Psicologia, uma literatura crescente sendo produzida desde 1943 e, no mundo de língua inglesa, desde 1960, sobre o auto-engano fatural, isto é, aquele que incide sobre a verdade ou falsidade de uma proposição. Este trabalho versará apenas sobre este tema, ou seja, nada será dito aqui a respeito do enganar-se a si próprio em relação a valores.

Na referida literatura o auto-enganado é descrito inicialmente - isto é, previamente a uma análise conceitual filosófica - como aquele que *acredita naquilo que ele sabe não ser verdadeiro*. A condição é em geral (mas nem sempre) vista como um caso de irracionalidade; mas, mais do que isso, a própria descrição acima parece ser irracional: ela não parece ser inteligível, já que ela soa fortemente paradoxal. De fato, como pode uma pessoa **acreditar** naquilo que ela **sabe** não ser verdadeiro? Na medida em que a pessoa está enganada, ela acreditaria que a proposição em questão

---

<sup>1</sup> Departamento de Filosofia – Faculdade de Filosofia e Ciências – UNESP – 17525-900 – Marília – SP. E-mail: trajano@marilia.unesp.br

é verdadeira; contudo, na medida em que ela é sabedora do fato de a proposição não ser verdadeira, ela não acreditaria (não poderia acreditar) que a proposição é verdadeira.

Com isso, o grande desafio para o filósofo nesse terreno tem sido o de produzir uma análise clarificatória, definicional ou não, que satisfaça os dois requisitos seguintes: (a) Capte o fenômeno em seu caráter genuíno de auto-mentira, isto é, não recorra à saída de tratar as expressões *auto-engano* e *auto-mentira* como uma tão-só maneira de falar para referir um certo tipo de erro cognitivo tão isento de paradoxo aparente quanto qualquer outro, e (b) Seja livre de paradoxos e satisfatoriamente inteligível. O presente trabalho é uma tentativa de apresentar uma tal análise.

Observo, de passagem, que este tema parece relacionar-se, e estreitamente, com o tema psicanalítico dos *mecanismos de defesa*, e com o tema marxista da ideologia como reflexo de interesses estranhos àqueles que ela alega (por exemplo, interesses econômicos de classe social). No entanto, não se tentará aqui estabelecer essas relações.

Ocorre que a atribuição de conhecimento e a imputação de crença ao auto-enganador, características da descrição paradoxal apresentada mais acima, são duas fontes independentes de dificuldades. No entanto, elas definem aquele que é o tratamento dominante na literatura. Referir-me-ei a ele como a a visão ou concepção **epistêmica** do auto-engano fatal.

Muitos teóricos do auto-engano têm sido incapazes de tornar inteligível uma explicação em termos de crença e igualmente incapazes de libertar-se dela. Nas mãos dos que insistem em caracterizar o auto-engano em termos de crença ou de crença-conhecimento, o fenômeno resulta eliminado de um modo ou de outro: aqueles que afirmam que o auto-engano é simplesmente crença motivada frente a evidência contrária, sem nenhuma implicação de má-fé, estão dizendo que a auto-mentira, estritamente falando, não existe, e aqueles que sustentam que o auto-engano é crença naquilo que o sujeito sabe ser falso estão dizendo (quer eles aceitem isso ou não) que a auto-mentira, estritamente falando, não pode existir. Em ambos os casos o fenômeno *evapora* à menção da expressão que visa descrevê-lo. Portanto, se desejamos manter o fenômeno diante de nós, parece que o que é preciso é o seguinte: nós teríamos de rejeitar não apenas a quadro (apresentado por Sartre, entre outros)

do auto-enganador como um possuidor de conhecimento claro, mas também a assunção, muito mais difundida, de que ele é um *acreditador* (de um certo tipo).

## 2 A definição e uma primeira argumentação em favor dela

Aquele que *cientemente* se engana a si próprio que uma sentença P é verdadeira frequentemente afirmará P, e seu ato de asserção não será do tipo que é característico da situação do mentir interpessoal: ele afirma P tanto para os outros quanto para si próprio. Além disso, ele não afirmará P com convicção completa, uma vez vez, estamos supondo, ele não está apenas em erro motivado mas também em *mauvaise-foi* (que é o termo de Sartre para a auto-mentira); noutras palavras, etc, no mínimo, terá algumas dúvidas sobre a verdade de P. Combinadas com a noção de disposição (para o comportamento verbal), as considerações acima sugerem a seguinte definição de auto-mentira fatural: *Um sujeito S está em auto-mentira fatural em relação a P no momento t se e somente se (1) S tem a disposição de afirmar P, (2) S tem, em algum grau, dúvidas sobre a verdade de P, e (3) O seu afirmar P não é um caso de mentira interpessoal propriamente dita.*

Segue-se agora o principal argumento para a introdução da noção de asserter na definição de auto-mentira.

Perguntemo-nos o que diferencia os casos problemáticos de auto-engano dos casos menos sérios de erro motivado. Sugiro que comecemos por dar a resposta a mais simples possível, mesmo que ela não seja muito esclarecedora nela mesma. Todos concordarão que isso que chamamos de insinceridade está, de um modo ou de outro, essencialmente presente. Este elemento é fartamente referido na literatura por meio de expressões que o exprimem ou o implicam: expressões como *erro deliberado, ignorância consciente, erro propositado, crença e descrença simultânea e consciente, mauvaise-foi, auto-mentira*, e também a fórmula batida: *O auto-enganador no fundo sabe que aquilo em que ele acredita é falso*. Essas expressões visam captar a condição que é essencial e específica ao auto-enganador; vou referir-nos a essa condição como sendo a da insinceridade-consigo-mesmo.

Podemos então dizer, grosso modo, que a auto-mentira é erro cognitivo motivado essencialmente acompanhado por insinceridade-consigo-mesmo.

Assumindo que a noção de erro cognitivo motivado não necessite de nenhuma clarificação especial para nossos propósitos, tudo o que temos que fazer, para obter uma definição apropriada do conceito, é clarificar a noção de insinceridade-consigo-mesmo e mostrar como, exatamente, ela se relaciona com o erro em questão; ou, noutras palavras, localizar com precisão o lugar da insinceridade-consigo-mesmo no erro motivado em questão.

A insinceridade, de modo geral, está envolvida em coisas tais como fingir, enganar, mentir. Portanto, em formas de comportamento. Quando somos sinceros ou insinceros, estamos engajados ou em expressar nossos sentimentos, preocupações, atitudes, verbalmente ou não-verbalmente, ou em comunicar nossas crenças. Desse modo, a insinceridade e seu oposto são eminentemente condições práticas; *sincero* e *insincero* aplicam-se a instâncias de **fazer**. Em particular, eles não se aplicam a estados mentais enquanto tais. O sentimento de insinceridade é um estado mental, sim, mas estados mentais, enquanto tais, não têm a propriedade de serem sinceros ou insinceros. Nem, igualmente, *estados da pessoa*.

Sim, é verdade que nós falamos como se eles tivessem essa propriedade. Dizemos, por exemplo, 'S1 está sinceramente envergonhado', 'A gratidão de S2 é sincera'. Mas perguntemo-nos qual é a análise correta dessas sentenças. É defensável, parece-me, analisar a segunda sentença como segue: Existe uma certa coisa X tal que (X é a expressão por S2 de sentimento de gratidão de S2 & X é sincera). De fato, se S2 não expressa de nenhum modo, nem mesmo para si próprio, seu possível sentimento de gratidão, como podemos aplicar *sincero* (ou igualmente *insincero*) neste caso? Analogamente, S1 está simplesmente envergonhado, e quando nós dizemos a primeira das sentenças acima, devemos ser interpretados como querendo dizer que sua mostra de vergonha, seja ela verbal ou não-verbal, é sincera. Do mesmo modo, dizer que a pessoa é sincera (ou insincera) é um modo de dizer que sua **manifestação** (de crença, de sentimentos, de atitudes) é geralmente, ou sempre, sincera (ou insincera).

Podemos ver isso mais claramente quando consideramos o seguinte. Contrariamente ao que ocorre com 'Ele insinceramente afirma que P', não faz sentido dizer 'Ele insinceramente acredita que P'. Não pode, portanto, fazer sentido estrito dizer 'Ele sinceramente acredita que P'. O que é transmitido por esta última sentença

é precisamente que a afirmação, P, da pessoa é uma expressão fiel de sua crença, isto é, é perfeitamente sincera. Estritamente falando, não existe o acreditar insincero. O mesmo se aplica, é claro, a 'Ele sinceramente pensa que P': não existe o pensar insincero que P. Igualmente, *sincero* e *insincero* não se aplicam a instâncias de saber. Acreditar-que-P e saber-que-P são estados mentais que, como tais, não são nem sinceros nem insinceros. Um estado mental de crença ou existe ou inexistente, mas ele não será sincero (ou *wholehearted*) ou insincero (ou *half-hearted*).

O mesmo se aplica ao tipo particular de insinceridade que está presente no auto-engano (e, igualmente, por sinal, no engano interpessoal). A insinceridade do auto-enganador sobre fatos só pode manifestar-se em uma instância de afirmação-de-fatos. Ser, de um modo ou de outro, insincero a respeito de uma crença fatural só pode tomar a forma de afirmar insinceramente a crença em questão. Em particular, aquele que mente para si mesmo a propósito de um fato, na medida em que ele responde a essa descrição, precisa dizer a si mesmo que a crença é verdadeira.

Este é um **insight** importante a respeito da ideia de sinceridade. Quando existe um certo tanto de insinceridade fatural para com os outros ou para si próprio, ela se manifesta em episódios de dizer, contar, assertar e semelhantes, e não nos estados de crença do agente. A pessoa é insincera ao comunicar sua crença ou conhecimento.

Podemos definir a sinceridade *fatural* dizendo que o agente é sincero quando, e apenas quando, ele acredita naquilo que asserta. Na verdade, assumindo a noção de asserção, então as idéias de ter crença fatural e a de sinceridade são interdefiníveis: podemos dizer que o agente tem a crença que P se e somente se o seu afirmar P é sincero. Tanto a idéia de asserção quanto a de acreditar estão pressupostas na definição acima de sinceridade. Mas o **insight** que estou tentando trazer à tona é que a sinceridade se relaciona com asserções de um modo no qual ela não se relaciona com crenças. E expliquei esse modo particular dizendo que os predicados *sincero* e *insincero* devem ser tomados como aplicando-se a **atos de asserção**, e não a **estados de crença**.

Elaboremos um pouco mais esse ponto. O auto-mentir está ligado com o uso da linguagem tanto quanto a mentira interpessoal está. E a linguagem em

questão é aquela na qual é possível afirmar fatos. Se não fosse pela dualidade de acreditar e dizer, inexistiria o mentir para si mesmo exceto num sentido totalmente metafórico. Se um ser é privado de uma linguagem que o capacite a afirmar fatos, como poderia ele mentir em algum sentido literal desse verbo? Por meio de acreditar conscientemente naquilo que ele sabe (portanto, acredita) ser falso? Mas, além de ser paradoxal, esta fórmula deixa completamente de captar a natureza do mentir ou da insinceridade, que é não o de um estado da mente mas de uma ação. O mentir, seja para os outros seja para si mesmo, não pode existir ou subsistir excepto alimentado pela dualidade do ato de asserção e da dimensão do acreditar. A sinceridade (sobre fatos) emerge com a possibilidade de assertar coisas outras que não aquelas nas quais acreditamos, e é definível apenas em termos do par crença-asserção. Quando o examinamos, esse ponto é óbvio, mas os teóricos do auto-engano surpreendentemente o perderam de vista.

Esta dualidade de acreditar e dizer está por trás de um dos pontos centrais apontados pelo filósofo Bernard Williams em seu artigo *Decidindo acreditar*, que é o seguinte: contrariamente ao que se sustenta habitualmente, há mais coisas na crença do que no conhecimento. Estou colocando o conceito de asserção no núcleo do conceito de auto-engano, e nosso conceito de asserção, escreve oportunamente Williams, “está vinculado à noção de decidir dizer algo que espelha ou deixa de espelhar aquilo em que voce acredita”. Sendo assim, é visível que a idéia de asserção deve ser um candidato muito bom para ajudar a iluminar o auto-engano, uma vez que aquilo que Williams está dizendo de ‘nosso conceito de asserção’ é, noutras palavras, que ele está atado à noção de sinceridade; e esta última é aquilo que especificamente está no centro do conceito de auto-engano. Pelo fato de a insinceridade (sobre fatos) estar tão intimamente atada quanto possível a assertar, este último é um candidato muito promissor para ajudar a captar o significado da auto-mentira.

É bom observamos que a insinceridade de que se trata aqui não é idêntica ao **sentimento** de insinceridade. Este último é um elemento caracteristicamente mental que acompanha algumas instâncias de insinceridade. mas ele não pode servir como um condição definidora da insinceridade. É bom que se mantenha presente este truísmo quando se está estudando o auto-engano, porque é

possível que o auto-engano sempre vá junto com algum sentimento de insinceridade e, sendo assim, poder-se-á ser tentado a dizer que, no auto-engano, a insinceridade é simplesmente aquilo que se sente quando se está enganando a si mesmo, e, segundo a visão que estou defendendo, é muito mais que isso: é uma propriedade que se aplica a ações linguísticas.

Assertar é, é claro, algo que fazemos, e um exemplo é o mentir interpessoal. E o que estou sugerindo é que uma outra ilustração, e uma que é particularmente interessante, é a auto-mentira. De fato, se sustentamos que a auto-mentira não tem nem um pouco a natureza de um ato, nós eliminamos toda a literalidade do termo, e nós não queremos isso; queremos reter um elemento genuíno de **mauvaise** foi propriamente dita. Noutras palavras, ao interpretar *mentira* em *auto-mentira*, queremos reter algo do sentido literal que o termo tem em *mentira para os outros*. Assim, se não definimos a auto-mentira em termos de ação - e , se definimos, o candidato óbvio, e talvez inevitável, é o ato de asserção - mas em outros termos, nós nos desligamos do fenômeno. 'Mentir' (para os outros) é um verbo de ação e não um que denota um estado da mente, e se, em seu uso reflexivo, seu caráter de ser um verbo de ação não é preservado, ele perde inteiramente seu sentido literal.

Para aqueles filósofos que estão realmente interessados no auto-engano - i.e., interessados o suficiente para concordar que alguma forma de insinceridade é da essência dele - alguma definição em termos de assertar é, quer eles estejam preparados para aceitar isso ou não, o único tipo de definição que pode ser satisfatório. Isso é assim, simplesmente porque a **mauvaise-foi** sartreana tem de ser entendida em termos de ser insincero, e não podemos ser insinceros (a respeito de fatos) a não ser por meio de/em/enquanto/através de, assertar/assertando algo.

É o auto-enganador alguém que acredita, um *acreditador* (de algum tipo)? De acordo com a presente análise, a resposta correta é que o auto-enganador deixa de, ou não consegue, acreditar naquilo que afirma. Enquanto alguém que mente (para si mesmo), ele justamente deixa de acreditar naquilo que ele está preparado para afirmar. A idéia de **mauvaise-foi** ou auto-mentira não combina com a de acreditar. Uma das coisas a respeito do auto-enganador que salta à vista é que, não importa quão insistentemente ele volte para, e consulte, a evidência favorável (ou aquilo que ele considera como tal), na medida em que ele pode ser descrito como



mentindo para si próprio, ele não consegue se tranquilizar a respeito da verdade da crença: ele não consegue convencer-se a si próprio. Podemos dizer - e esse modo de dizer torna o quadro convenientemente dramático - que toda e qualquer tentativa que ele faz para convencer-se da verdade daquilo que ele diz para os outros e para si mesmo, acaba em fracasso.

O exame cuidadoso tanto do significado de *mauvaise-foi* quanto de nossas percepções intuitivas sobre o que *rola* quando uma pessoa é descrita como estando mentindo para si mesma, dá suporte para a visão que está sendo defendida aqui. Quando nós *examinamos* o auto-enganador, o que vemos é uma pessoa que asserta algo, e não uma pessoa que acredita no que asserta. *Auto-mentira* denota não tanto um estado da mente quanto um conjunto de atos verbais - ou, mais exatamente, a disposição para executar tais atos - que são constituídos por episódios individuais de a pessoa alegadamente estar mentindo para si mesma. E esses episódios são episódios de assertar. Aquele que mente para si próprio é um **assertador**, e um assertador que não é um **acreditador**.

Claro, aquele que mente para si mesmo a respeito de P alegaria não ter dúvidas sobre P; ele diria ser perfeitamente sincero a assertar P, e sustentaria ser um acreditador. Podemos dizer que é esta alegação implícita de acreditar, acoplada com o persistente fracasso em acreditar, que faz dele um auto-enganador. A auto-mentira é um tipo especial e único de assertar-sem-acreditar, a saber, aquele no qual a pessoa alega para si mesma, e não apenas para os outros, que ela acredita naquilo que diz; enquanto o mentiroso (normalmente) descrê daquilo que ele diz e está ciente disso, o auto-enganador deixa de acreditar naquilo que ele asserta, e não vê isso claramente.

Consequentemente, o auto-enganador não é, estritamente falando, um **auto-enganado**. Pode-se dizer que ele tenta enganar-se, mas não há na auto-mentira nenhum estado de engano bem-sucedido. Aquele que mente para si mesmo sobre P, na medida em que ele responde a esta descrição, não tem êxito em enganar-se a si próprio que P é verdadeiro. No momento em que ele tivesse êxito, ele passaria a ter verdadeiramente a crença que P com completa convicção e, desse modo, ficaria livre de *mauvaise-foi*. A maioria dos filósofos se concentrou sobretudo na conotação do substantivo *auto-engano*, ao invés de na de *mentir para si mesmo*. Eles assumiram que

a dualidade, existente na situação interpessoal, do processo (isto é, da conduta) de enganar e do produto dele (o resultante estado de engano) vige de algum modo também na modalidade reflexiva: como resultado de embarcar nas manobras auto-enganadoras, o sujeito terminaria por colocar-se a si mesmo no estado de auto-engano. Esses filósofos acabaram enclacrados na idéia de que um tal paralclismo vige. Mas, de acordo com a presente análise, na situação de mentira reflexiva o estado de engano (estritamente: estar enganado que P) é inexistente. Focar a questão ‘Como é estar em estado de auto-engano?’ não é um bom procedimento, porque ele fortalece a inclinação de assumir que, no que diz respeito a erros cognitivos motivados, há sempre um estado de engano (de novo, especificamente: estar enganado que P) no qual o sujeito está. É melhor começar com a consideração de *enganar-se a si mesmo*, a qual conduz mais ou menos diretamente para o interior da análise de *mentir-para-si-mesmo*. Quando este último, por sua vez, é adequadamente analisado, nós podemos garantir que não há como postular um estado de engano. Mas precisamos enfatizar, por outro lado, que o auto-enganador, a despeito de não estar enganado que P, está contudo em erro cognitivo a respeito de P. Um erro sério, reconhecidamente.

Mas, por outro lado, é igualmente importante chamar a atenção para as similaridades entre o acreditador e o auto-enganador. Este último assera P ao mesmo tempo em que tem para com a crença que P uma atitude que é, em aspectos importantes, muito semelhante àquela **mantida** por alguém que tem a crença em questão. Primeiro, aquilo que ele diz para nós ele diz também para si mesmo; segundo, ele está preparado para exhibir, sem fingimento, um padrão possivelmente complexo de comportamento – isto é, ações de vários tipos - que é bastante próximo daquele exibido por alguém que acredita na verdade de P. Nós, observadores, podemos mesmo ser tentados, com o fim de explicar muitas daquelas ações, a atribuir a ele a crença em P. Estes dois aspectos fazem dar-nos conta de que o auto-enganador está bastante perto do acreditador e distante do mentiroso. O fato, contudo, de que há analogias visíveis e importantes entre eles não constitui, em si mesmo, uma objeção para a visão de que o acreditar está ausente da auto-mentira. E devemos resistir à tentação de dizer que, por força das analogias, o auto-enganador **acredita** - num modo todo especial de acreditar - naquilo que ele afirma. Pois é importante insistir em que o auto-enganador é um assertador que não é um acreditador, e com isso enfatizar seu fracasso em acreditar.

Ocorre uma outra coisa com o auto-enganador que dá algum apoio à visão que está sendo proposta aqui. O auto-enganador é algumas vezes, e corretamente, descrito como alguém que sustenta *enfaticamente demais* ou *alto demais* - e, podemos acrescentar, insistentemente demais - aquilo que ele alega ser verdadeiro. Este aspecto é talvez especialmente saliente quando o assunto com respeito ao qual ele está mentindo para si mesmo é central para sua vida e é sentido por ele como particularmente importante. Podemos presumir, neste caso, que, em virtude de sua centralidade e importância, o assunto permanecerá recorrente para ele, e que ele frequentemente estará repetindo (tanto para os outros quanto para si próprio) que as coisas são como ele alega que são. Ele permanecerá dizendo para si próprio e para os outros que P, porque as ocasiões que apresentam um desafio à sua crença não cessarão, presumivelmente, de confrontá-lo.

Este aspecto do auto-enganador se encaixa bem na análise que está sendo defendida aqui. Este é, claramente, um aspecto em conexão com o qual a natureza linguístico-assertórica que estou imputando à auto-mentira sobressai-se bastante vividamente. De fato, se é importante para a pessoa que P seja verdadeira, então, em virtude das dúvidas que ele estará todo o tempo experimentando, é compreensível que ele tenha uma necessidade forte e persistente de tranquilizar-se a si próprio que a crença é bem-fundada. Se você tem dúvidas que você não admite como tais, então a própria existência delas torna necessário que você se tranquilize a si próprio, e isto você fará agarrando-se a quase qualquer consideração ou evidência favorável de que disponha. Em contraste vivo com isso, quando você está completamente convencido da verdade de P, você normalmente não tem nenhuma necessidade particular de ficar dizendo para você mesmo, de tranquilizar-se, que P é verdadeira. Você se lembrará, ou será lembrado, que P e/ou assertará P para os outros ou para você próprio quando a ocasião apropriada surgir. Para focar o ponto mais nitidamente, imagine que você acredita acertadamente e completamente, sem a menor sombra de dúvida - nem mesmo uma reprimida no inconsciente ou disfarçada em alguma outra coisa por seus *mecanismos de defesa* - que você é extremamente talentoso. Imagine, além disso, que frequentemente você pare para contemplar o fato de você ser muitíssimo talentoso; nestas ocasiões você poderá, por exemplo, considerar um ou outro aspecto desfrutável dessa sua feliz condição. Mas, uma vez que, por hipótese, você não está em erro (motivado), seria incorreto dizer que você

fica dizendo/afirmando para você próprio que de fato você é muito talentoso. Você não estará fazendo isso, porque, entre outras coisas, você não precisa fazê-lo. *Este é o modo de ser das crenças inteiras, impregnadas por inteiro de convicção, e é o exato oposto daquilo que está envolvido na auto-mentira, a qual é justamente caracterizada por (um tipo único de) meia-convicção.* Podemos dizer, num modo convenientemente compactado: pelo fato de o auto-enganador não conseguir acreditar, ele tem uma necessidade característica e constante de dizer a si próprio que a crença é verdadeira. No caso extremo podemos presumir que ele estará obcecado com contar para si mesmo a estória que ele deseja ser verdadeira, e nesta obsessão podemos convenientemente discernir a verdadeira face do auto-enganador: aquela de um assertador cujo comportamento verbal trai seu fracasso em acreditar. Vemos, desse modo, que o recurso à linguagem do fazer-asserções, longe de ser um modo fabricado de evitar a idéia problemática de ter-crenças, justifica-se à luz de alguns fatos presumidos a respeito do mentir reflexivo. A proposta que estou defendendo tem êxito em captar o aspecto do auto-engano em questão.

Mas tem mais. A presente análise se conforma bem com algumas *acusações* que, algumas vezes, estamos dispostos a fazer ao auto-enganador. É possível, por exemplo, fazer a este a seguinte acusação: ‘Você só vê aquilo que lhe convém’. Quando esta sentença é tomada como uma crítica, e não como, digamos, meramente uma advertência para re-pensar a coisa, o que ela significa é não que a pessoa realmente vê as coisas que, por alguma coincidência, são aquelas que lhe são favoráveis, e não vê nada mais; mas, antes, que ela **afirma/alega ver** apenas as coisas favoráveis. E alegar é assertar, não é acreditar. Uma outra **observação** que as vezes dirigimos ao auto-enganador, e que também está em conformidade com a análise aqui proposta, é ‘Você não acredita realmente no que você está dizendo’. Nós não podemos objetar, com razão, desses modos a alguém em auto-engano ou simplesmente em erro.

Outro ponto. A **mauvaise foi** que se presume envolvida no auto-engano dá algum suporte tanto para a visão segundo a qual o auto-enganador é as vezes moralmente condenável, quanto para a presunção da experiência, por ele, do sentimento de culpa. Esta presunção está em harmonia com a definição, de natureza conativa, que está sendo adovagada aqui. De fato, não é normal sentir-se culpado, nem ser moralmente condenado, por exemplo, por **acreditar** em algo; normalmente

sente-se culpa, ou se é condenado, por alguma coisa que se **faz**. O que o auto-enganador faz é assertar P, e isso ele sente, ainda que confusamente, não estar justificado em fazer, uma vez que tem dúvidas a respeito da crença. Ele, de algum modo, percebe que está agindo de um modo indevido.

### 3 Segundo momento da argumentação

Os teóricos que levam a sério o auto-engano – isto é, aqueles que interpretam *auto-mentira* literalmente - sempre acharam necessário, e com razão, é claro, imputar ao auto-enganador algum tipo de *conhecimento* da verdade da qual ele tenta fugir. Segundo uma das fórmulas preferidas, 'Bem no fundo ele sabe da verdade'. Este *conhecimento* é considerado um aspecto distintivo importante daquele que mente para si próprio, em contraste com aquele que incorre em outros erros motivados. Na definição que está sendo defendida aqui, este aspecto é formulado como aquele de ter dúvidas a respeito da verdade da crença correspondente.

Poder-se-ia, no espírito dos argumentos envolvendo paradoxos, levantar a seguinte objeção, aliás bastante óbvia, à definição oferecida: também na definição proposta, infelizmente o fenômeno não pode ocorrer, já que nenhuma pessoa sã seriamente afirmaria P tendo dúvidas sobre sua verdade, a menos que ela esteja apenas mentindo para os outros. A direção da réplica talvez esteja visível: ela não vê claramente que está em dúvida. Ter dúvidas significa aqui sentir-se inseguro a respeito da verdade de P, e o auto-enganador não vê claramente que ele está inseguro - ou, talvez melhor, que ele está inseguro a respeito da verdade de P. A sugestão é que ele sente isso **confusamente**. Ele tem uma **percepção confusa** do que ele está fazendo, um sentimento brumoso de que ele carece de boas razões, ou tem boas razões contrárias, para afirmar a crença. Em particular, e se a ocasião aparecer, ele negará que as dúvidas experimentadas por ele são de fato dúvidas. Ele tentará explicá-las como sendo outra coisa; ele alegará crer realmente que P. E nessas negações e alegações ele estará incorrendo em erro. Mas este não será um erro independente do outro; ele se reduz ao mesmo erro/mentira original: negando, ou renegando, as dúvidas ou alegando acreditar que P, ele estará ignorando as dúvidas que, no entanto, estará sentindo.

Esta sua negação não o força necessariamente a negar que ele

experiencia algum sentimento ruim em conexão com P, mas ela impede que ele reconheça a perturbação como insinceridade. Não é difícil explicar o sentimento ruim que se tem de um modo que satisfaça os próprios desejos: 'É que a gente gostaria de ter certeza absoluta, e isso simplesmente não é possível em assuntos como esse', 'É falta de confiança devida a cansaço/ansiedade/etc'. Para que a negação em questão seja possível, é preciso que o sujeito não esteja implicitamente ciente da verdadeira natureza do seu sentimento ruim. Desse modo, vemos que na auto-mentira a atitude do sujeito para com seu próprio estado de crença é semelhante a, e tão importante quanto, sua atitude, por exemplo, para com a evidência desconfirmatória forte (quando esta existe): ele afirma/alega sem realmente acreditar.

Por outro lado, precisamos enfatizar que, embora as dúvidas precisem ser, até certo ponto, mascaradas como uma outra coisa, elas necessitam, como foi indicado atrás, serem de **algum modo sentidas como dúvidas**. Dúvidas, estritamente falando, são uma condição que é experimentada pelo sujeito. Não importa quão negadas e deixadas num estado mal-definido, elas são, ao menos parcialmente, uma experiência mental do sujeito. Elas existem como algo que é subjetivamente experienciado. A percepção intuitiva que temos do auto-enganador inclui um elemento de confiança quebrada, de inquietação ou perturbação. Na medida em que a pessoa está mentindo para si mesma, sua confiança na verdade da crença permanecerá abalada, e esta confiança abalada, junto com a perturbação, provém precisamente do sentimento de dúvida (não-admitida).

Dizer que o auto-enganador sente ou experimenta dúvidas significa atribuir a ele algum modo de ciência ou consciência. E, é claro, uma questão crucial que qualquer clarificação completa do auto-engano deve responder é 'Qual é a exata natureza dessa consciência?' Sabemos que ela deve ser tal que seja compatível com a continuada asserção da crença pelo sujeito. Não farei aqui nenhuma tentativa de chegar a uma grande clarificação desse ponto difícil. Mas farei algumas observações com o fim de elucidá-lo numa certa medida. Penso que ajuda dizer que o auto-enganador está num certo estado de **estar-confusamente-ciente-de**. Ao dizer isso não estou pensando, por exemplo, num estado de confusão admitida com respeito à natureza de seu próprio sentimento. Não ocorre que ele esteja perfeitamente ciente de que seu sentimento é um sentimento mal-definido. Estou pensando que o estar-

ciente é ele próprio confuso: na medida em que ele sente incerteza, ele pode, num certo sentido, ser descrito como estando ciente dessa incerteza, mas na medida em que ele nega o sentimento, seu estar-ciente precisa ser um estar-ciente **confuso**. Dado que a pessoa afirma sem mentir para os outros, segue-se que as dúvidas mal-definidas das quais ele está confusamente ciente são aquilo que define o auto-engano. Na medida em que ficamos dentro de uma proposta apenas definicional do auto-engano fatural, é exatamente a possibilidade de negar as dúvidas que são, não obstante, experimentadas, que é importante e mesmo crucial.

Não parece haver nenhum modo consistente de surpreender o elemento de **mauvaise foi** exceto postulando um sentimento de dúvida. Este último, junto com sua negação, são aquilo que constitui a pessoa em erro motivado como uma pessoa que é insincera consigo mesma. A insinceridade consigo mesmo, tanto como uma experiência privada do sujeito quanto como algo manifesto para os outros em seu comportamento observável, é uma indicação valiosa de erro cognitivo motivado e, portanto, uma ajuda em direção a superar este último. O sentimento de dúvida, conseqüentemente, fornece parte da base para superar o auto-engano. Não surpreende, assim, que esse sentimento acabe emergindo como um coisa *boa*: a parte do sujeito que tem dúvida, nós poderíamos dizer, é a parte racional do sujeito. Com efeito, embora a auto-mentira seja algo irracional, há nela algo que é profundamente racional, a saber, as dúvidas que existem na mente do sujeito e agem como um obstáculo poderoso contra a real adoção por ele da crença sem fundamento ou mal-fundada. A superação da auto-mentira é a prevalência das dúvidas sobre sua negação.

Pelo fato de o auto-enganador não poder encontrar um modo satisfatório de se livrar da evidência adversa, ele adota a política (para emprestar uma expressão de Herbert Fingarette) de evitar um exame cuidadoso desta última. Ele evita confrontar-se frente à frente com a evidência perturbadora, porque, se ele fosse considerá-la atentamente, ele deixaria de ser capaz de reter sua crença. Analogamente, uma consideração cuidadosa das razões para sua alegação revelaria seu sentimento mal definido como sendo o de ter dúvidas. Sendo assim, para manter-se no erro, ele precisa evitar que o referido sentimento se torne bem-definido. De modo geral, a auto-mentira floresce num ambiente de confusão e de falta de clareza, e isto se mostra de modo particularmente visível na atitude do auto-enganador para

com seu sentimento de dúvida: ele prefere deixar este último brumoso e resistirá a tentativas de clarificá-lo.

Reflitamos um pouco mais sobre a natureza da *ciência* que o auto-enganador tem de suas dúvidas. A idéia de estar confusamente ciente de dúvidas é a de ter dúvidas e negá-las, e não parece haver nada de paradoxal nisso. Estamos bastante familiarizados com esse tipo de situação: tenho ciúmes de alguém ou estou muito decepcionado com algo, com isso experimento o correspondente sentimento desagradável, e inclusive ajo e reajo exatamente como uma pessoa com ciúmes ou decepcionada faria. Contudo, recuso-me a admitir, não apenas para os outros mas também para mim mesmo, que estou com ciúmes ou decepcionado. Posso admitir meu sentimento desagradável enquanto tal, mas então darei a ele alguma outra interpretação; e, usualmente, não é difícil encontrar bons candidatos para isso. Mais tarde poderei confessar que estava realmente com ciúmes ou desapontado e que eu 'no fundo estava ciente disso'. Pois bem, o mesmo parece ser verdadeiro a respeito de dúvidas negadas: posso explicar o fato de estar me sentido incerto a respeito de P numa determinada ocasião, dizendo que a insegurança que experimento é devida, por exemplo, à conduta rude de meu interlocutor, que fica me intimidando, e não ao fato de não estar convencido da verdade de P. E mais tarde eu poderei rever e corrigir essa alegação. Desse modo, no que diz respeito à ciência que o auto-enganador tem *da verdade* (em minha análise *das dúvidas*), a analogia plausível é com ciúme/decepção negado e coisas afins. Além disso, não há nada implausível na idéia de que o estar-confusamente-ciente é capaz de guiar os passos do sujeito numa direção ou noutra.

Dizer isso não significa dizer, é claro, que dúvidas (em geral, estados mentais) existentes e negados não sejam bastante intrigantes e difíceis de explicar. O problema de dar uma explicação para a ocorrência de tais fenômenos tem que ser respondido por uma teoria explanatória psicológica, não por uma mera análise filosófica definicional. Mas não precisamos de um tal teoria para ter uma definição, para propósitos operatórios, do auto-engano fatural. O presente estudo da auto-mentira fatural está dirigido apenas para o fim de chegar-se a uma definição razoavelmente aceitável dela, com vistas a produzir clarificação conceitual.

Tentativas filosóficas de definir a auto-mentira invariavelmente recorrem a alguma distinção entre formas de consciência ou de estar-ciente. A presente



tentativa não é exceção. A distinção que está por detrás da definição em termos de assertar-ao-mesmo-tempo-que-se-duvida é a distinção entre consciência confusa e consciência clara dos próprios estados mentais.

#### **4 Terceiro momento da argumentação**

Esta seção final consiste de uma série de considerações que visam mostrar que a análise proposta satisfaz um certo número de requisitos que seria justificável exigir de qualquer definição de auto-engano fatual.

O primeiro ponto, embora óbvio, merece ser mencionado. Qualquer fórmula que queira definir a auto-mentira precisa fazer uma referência essencial a crenças fatuais. A apesar de ter deixado de fora de minha explicação a idéia de acreditar (numa proposição fatual), ainda assim retenho a condição de crenças fatuais estarem essencialmente envolvidas na auto-mentira, uma vez que para cada instância de fazer-asserções, existe uma correspondente crença fatual que é precisamente aquilo que é assertado. Desse modo, minha proposta não deixa de satisfazer esse requisito.

O segundo ponto é o importante e conhecido requisito de que a definição faça justiça à nossa visão intuitiva segundo a qual algum tipo de mentira é elemento essencial do quadro. A análise proposta identifica a auto-mentira como uma instância de assertar. A idéia geral de insinceridade ou mentir (fatural) é a de dizer algo diferente do que se acredita. Portanto, o fenômeno do auto-engano, quando descrito nos termos propostos, de fato inclui a insinceridade, já que naqueles termos existe uma lacuna - mesmo que seja o tipo menor possível de lacuna - entre aquilo que é assertado e aquilo que é acreditado: embora negadas, as dúvidas são reais e são prova de uma falta de acreditar.

A maioria, se não a totalidade, dos teóricos do erro cognitivo motivado concorda que o *problema da consciência* – isto é, o problema do que é estar consciente de algo e de quais são as variedades de estar-consciente-de - está no centro do estudo do fenômeno. Se é assim, qualquer clarificação conceitual aceitável dele não pode deixar de levantar esse problema. E, pelo fato de a noção de estar-confusamente-ciente-de subjazer à definição proposta, esta última está em consonância com aquela suposição. Na verdade, a noção específica de dúvidas-que-

são-ao-mesmo-tempo-experimentadas-e-negadas coloca o problema da consciência num foco particularmente nítido.

Uma definição de mentira reflexiva precisa, naturalmente, diferenciá-la claramente da mentira interpessoal. E muito do sentido, tanto no uso pré-analítico quanto em formulações filosóficas, de descrever o auto-enganador como um acreditador é precisamente com o fim de fazer essa distinção. Talvez se tenha notado que, na análise proposta, aquela distinção é feita simplesmente por meio da estipulação explícita de que o assertar não é um caso de mentira interpessoal. Parte do que fiz, ao formular a definição proposta, foi substituir **acreditar** por **assertar-para-si-próprio**. Acreditar fica, desse modo, analisado *para fora* (isto é, eliminado), sem detrimento do papel que ele desempenhava, o qual é retido e assumido por assertar.

Um quinto requisito é o seguinte. O auto-engano é descrito, e corretamente, como constituindo uma *condição inerentemente conflituosa*. E, de fato, nenhuma definição adequada dele pode deixar de implicar, ou ao menos sugerir fortemente, que a situação de *conflito interno* vige. Não é difícil ver que a análise proposta se sai bem nesse teste. O conflito é aqui retratado num ato de asserção que é acompanhado por dúvidas que, embora negadas, são sentidas, e as vezes intensamente, como dúvidas. Por causa dessas últimas, o ato de asserção do auto-enganador é algo que ele sente não estar muito justificado em fazer. Existe uma tensão, podemos dizer, entre a parte dele que tem dúvidas e a parte dele que faz a asserção. O assertar não pode coexistir harmoniosamente com o ter dúvidas, um está em conflito com outro, e isso é inerente à condição da auto-mentira. Aquele que mente para si mesmo não é capaz nem de calar a voz da dúvida, nem de desistir da afirmação de que a crença é verdadeira. Observei antes que é importante sublinhar o fracasso do auto-enganador em acreditar, e uma das razões por que isso é assim, é que neste fracasso reside uma dimensão do doloroso dessa condição: o que torna a condição do auto-enganador penosa é seu fracasso em acreditar naquilo que ele seriamente afirma. É em grande parte por isso que o auto-mentir é um *estado conflituoso*: se o auto-enganador de fato acreditasse naquilo que ele afirma ser verdadeiro, nenhum conflito interno seria experimentado. Pelo contrário, ele estaria satisfeito, já que acreditaria naquilo que ele prefere ser verdadeiro. O fracasso de acreditar condensa/exibe a natureza conflitiva do auto-engano.

O conflito em questão pode corretamente ser interpretado como um caso especial do conflito entre razão e sensibilidade. De fato, as dúvidas do auto-enganador são, podemos dizer, a voz da razão que fala dentro dele; por outro lado, o seu assertar P e sua alegação implícita de acreditar em P constituem a manifestação de seus desejos e quereres, isto é, sua paixão. E, de fato, a superação da condição de auto-mentira, portanto o fim da conduta insincera, é a vitória de dúvidas legítimas, portanto da razão.

Esse conflito pode ser mais bem explicado com a ajuda de uma certa noção de **necessidade de justificação cognitiva**. A luta do auto-enganador para persuadir-se a si próprio da verdade da crença reflete a importância da auto-justificação cognitiva. O conflito não-resolvido continua, porque o sujeito sente que a justificação que ele dá não é suficientemente boa; esse conflito reside no fato de que a auto-mentira pressupõe, por parte do sujeito, a aceitação de algo que ele, ao mesmo tempo, viola. A ocorrência da auto-mentira manifesta tanto a necessidade quanto a pobreza da justificação buscada. Como diz A. E. Murphy em seu livro *The theory of practical reason*, "... o auto-engano seria sem propósito (...) se a auto-justificação não fosse importante"; depois de escrever que "a 'racionalização' é a homenagem que o irracional presta à razão", ele acrescenta que "quando a ignorância é propositada, esta homenagem é, naturalmente, insincera. (...) Aqui, de fato, o mal é 'da vontade', é mal voluntário. Mas esta não é uma vontade que nada tem a ver com a razão. É uma vontade que precisa justificação, mesmo que ela necessite corromper o entendimento para garanti-la". Este tema é ecoado no livro *Self-deception*, de H. Fingarette, que afirma que "o entrar em auto-engano é motivado por uma preocupação com a integridade pessoal". A preservação da integridade pessoal significa, neste caso, satisfazer a exigência de justificação. Para que um real conflito interno emergja, temos obviamente de supor que a pessoa julga valiosa a preservação de sua integridade intelectual.

Procurei mostrar que a definição proposta faz justiça a traços importantes que aceitamos serem essenciais ao auto-engano, e que nessa medida ela é bem-sucedida em retratar o fenômeno genuíno. Além disso, ela não sucumbe vítima dos paradoxos familiares. Não há inconsistência no fato de o ato de asserção ser acompanhado de dúvidas não admitidas como tais, a menos que acrescentemos a

presunção de que o sujeito tem, de fato, a crença. Os paradoxos da auto-engano provêm ou de se presumir que aquele que mente para si próprio acredita no que ele alega, ou de se atribuir a ele conhecimento claro da verdade. E, no presente tratamento, não se está fazendo nenhuma dessas duas coisas.

O auto-engano é certamente um fenômeno intrigante.: as idéias de dúvidas- negadas e de estar-confusamente-ciente-de necessitam de mais clarificação e explicação. Contudo, esta não é uma tarefa de definições filosóficas (clarificatórias), e sim de uma teoria científica explanatória da consciência. Falta de clareza é uma coisa, paradoxo é bem outra. Qualquer que seja a falta de clareza da definição proposta, proveniente da idéia de dúvidas negadas das quais o sujeito está confusamente ciente, ela não parece ser devida à definição propriamente dita; ela provém, antes, de nossa ignorância a respeito dos meandros e detalhes da (in)consciência humana. Na verdade, a definição proposta parece ter a vantagem de focar nitidamente o problema explanatório do auto-engano.

Para concluir. Os propósitos legítimos que uma explicação em termos de crença/conhecimento – isto é, a epistêmica – procura preencher, são preenchidos pela definição proposta. Portanto, não parece haver nenhum aspecto com respeito ao qual o tratamento epistêmico seja superior ao aqui defendido. Este último determina o caso de erro motivado como caso de auto-engano, e, além disso, ele retém o fenômeno em sua inteireza (e também, é verdade, em seu caráter intrigante). Em análises filosóficas estamos melhor sem fórmulas paradoxais como “Ele (de algum modo) acredita no que ele sabe ser falso”. Portanto, não devemos usá-las, a menos que precisemos. E, se a presente análise for julgada aceitável, não precisamos.

### **Referências Bibliográficas**

- FINGARETTE, H. *Self-deception*. London: Routled & Kegan Paul, 1969.
- MURPHY, A. E.. *The theory of practical reason*. La Salle: Open Court, 1964.