

Ceticismo e Reconhecimento
Outras Mentes, Racionalidade e Linguagem

Fernando Matias Valenzuela Carlucci

Orientado por Jônadas Techio

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

Porto Alegre, Brasil

CIP - Catalogação na Publicação

Carlucci, Fernando
Ceticismo e Reconhecimento / Fernando Carlucci. --
2018.
150 f.
Orientador: Jonadas Techio.

Tese (Doutorado) -- Universidade Federal do Rio
Grande do Sul, Instituto de Filosofia e Ciências
Humanas, Programa de Pós-Graduação em Filosofia, Porto
Alegre, BR-RS, 2018.

1. Filosofia. 2. Ceticismo. 3. Reconhecimento. 4.
Wittgenstein. 5. Cavell. I. Techio, Jonadas, orient.
II. Título.

Fernando Matias Valenzuela Carlucci

Ceticismo e Reconhecimento
Outras Mentes, Racionalidade e Linguagem

Tese apresentada ao Programa de Pós-Graduação de
Filosofia da Universidade Federal do Rio Grande do
Sul como requisito parcial para obtenção do título de
Doutor em Filosofia

Banca Examinadora

Prof. Doutor Paulo Estrela Faria (UFRGS)

Prof. Doutor Eduardo Medeiros Vicentini (UFSM)

Prof. Doutor Eros Carvalho (UFRGS)

Prof. Doutor Victor Krebs (PUC Perú)

Porto Alegre, 2018

Para Antônio e Loreto

Resumo

Esta tese de doutorado é uma investigação do conceito de reconhecimento e das suas conexões com o problema do conhecimento de outras mentes e de outras vontades. Este conceito é oferecido como uma forma de dissolver o problema cético de outras mentes, mas na mesma medida em que ele dissolve o problema, tal conceito abre um rico campo de exploração filosófica.

O conceito de reconhecimento resulta do fracasso das tentativas de atender a demanda cética por certeza da existência de outras mentes. O fracasso desta empreitada epistemológica implica um recuo onde se passa a analisar as bases da significatividade do conceito de mente. Uma vez analisadas as condições de sua significatividade, o problema cético será compreendido na sua origem e passarei a investigar os diferentes modos de relação com o outro. O resultado será a possibilidade de conhecimento de outras mentes, não como função do conceito de certeza, mas como forma do reconhecimento do outro. A partir de então passo a investigar a gramática do fenômeno de reconhecimento — sua unidade, a lógica pronominal que a compõe, as formas autênticas de reconhecimento por contraste com as formas projetivas do eu e finalmente a conexão entre conhecer outras mentes e o reconhecimento do outro como um ser dotado de vontade.

Palavras-chave: ceticismo; disjuntivismo epistemológico, reconhecimento, ética do reconhecimento, normatividade, auto-conhecimento, racionalidade.

Abstract

This dissertation investigates the concept of recognition and its connections with the problem of other minds and other wills. This concept is presented as a way to dissolve the skeptical problem of other minds, and as a further consequence it also opens a rich field of philosophical investigation.

The concept of recognition is first grasped in the responses to the skeptic's demand for certainty about the existence of other minds. The modern epistemological enterprise's failure convinced philosophers to step back from the original problem and motivated a new posture in the analysis of the concept of mind. Once this task was accomplished, the skeptical problem is understood in its origins and this enables the investigation of the different modes of relation to the other. The result will be the possibility of knowledge of other minds, however not as a function of the concept of certainty, but as a form of recognition of other minds. This dissertation will also seek the understanding of the grammar of recognition — its unity, the pronouns that constitute its logic, the authentic forms of recognition and finally the connection between knowing other minds and the acknowledgment of the other as a being gifted with his/her own will.

Keywords. Skepticism, epistemologic disjunctivism, acknowledgment, ethics of recognition, normatively, self-knowledge, rationality.

Table of Contents

Agradecimentos 7

Introdução 8

Chapter 1 - The Problem of Other Minds 17

1. From me to others 17
2. The (not so) old problem of other minds 18
 3. Arguments from Analogy 20
 - 3.1. Russell's Knowledge by Description 21
 - 3.2. Two other attempts 23
4. The conceptual Problem of Other Minds 25
5. Kripke and the Rule-Following Problem 26
 - 5.1. Three shots against Skepticism 28
 - 5.2. The skeptical solution 30
6. Wright and the Fact of Meaning 32
 - 6.1. Wright's Constitutivism 34
7. McDowell and Disjunctivism 36

Chapter 2 - From the Concept of Other Minds to Acknowledgment 45

1. Summary 45
2. Theory Theory 46
3. Interaction Theory 48
4. Strawson and Criteria for Others 49
5. Nagel's Sartre and Look of the Other 53
6. Cavell's Knowing and Acknowledging 55

Chapter 3 - The Epistemology of Recognition 70

1. The Epistemology of Recognition 70
2. Honneth: Recognition as Social Affirmation 73
3. Margalit and Asymmetric Recognition 78
4. Thomas Reid and the Idea of Social Acts of Mind 82
5. Peacocke: The logic of the Second Person 83
6. Rödl: Recognition as Joint Activity 88
 7. Clearing things up 94
8. Haddock: Between Passivity and Activity 96
9. Humans Condemned to Expressiveness 100

Chapter 4 - The Ethics of Recognition 104

1. Summary 104
2. Dignity and Rationality 105
3. Darwall and Second Person Standpoint 110
 - 3.1. A reading of Kant's Categorical Imperative 110
 - 3.2. Moral Obligation and Address 113
 - 3.3. Agent-relative and Agent-neutral 115
 - 3.4. Second Person and Moral Community 117
 - 3.5. Derivation of Moral Principle 118
 - 3.6. Objections 122
 4. Korsgaard's Alternative 124
 - 4.1. Self-Consciousness 124
 - 4.2. Self-Constitution Argument 125
 - 4.3. Interaction Argument 129
 - 4.4. Some Remarks and Objections 132
5. Final Remarks 134

Conclusão 144

Agradecimentos

Em uma tese em que o conceito de reconhecimento é debatido, talvez esta será a única vez em que falarei sobre uma forma específica de reconhecimento — o agradecimento. Agradecer é uma forma de reconhecer os esforços de outras pessoas em projetos que não são fins últimos em suas vidas. A vida humana se fez de modo que é possível olhar para o lado com ternura a fim de doar um pouco do seu tempo a outros, uma vez que todos, em algum momento, precisamos de ajuda. Aqueles que são ajudados tem o dever de sentirem-se gratos e por isso, hoje orgulhoso, cumpro meu dever e agradeço.

Agradeço àqueles que chamo de pequena e grande família. Meus pais, Antônio e Loreto, pelo amor e pela constante compreensão. Sem eles nada disso seria possível. Agradeço ao que chamo de grande família composta pelos amigos que encontrei nesta trajetória. Aos que sempre permitiram que eu expressasse minhas angustias e que sempre entenderam minhas palavras mesmo quando era difícil a mim mesmo pronunciá-las: Adenauer, Cássio, Greice, Ísis e Karina. Agradeço aos amigos que amo e que sempre me acolheram com sinceridade e tolerância: José Eduardo Porcher e Magda Togni.

Agradeço ao meu orientador Jônadas Techio. Professor e amigo que me ensinou não somente a entender o que eu sempre mais desejei entender em filosofia, mas me ensinou a pensar em liberdade e a ter um temperamento filosófico.

Agradeço à Deborah Fuzina, mulher que amo e que dá sentido aos meus dias.

Agradeço aos amigos Giovanni Rolla e Rafael Ribeiro com quem pude rir e aprender filosofia durante estes quatro anos, ao professor Eros cuja seriedade tanto admiro, aos professores Raphael Zillig e Silvia Altmann que apostaram em mim lá no início de tudo. Agradeço ao pessoal da secretaria do PPG em Filosofia que tanto me ajudou em tramitações burocráticas — as famosas papeladas.

Agradeço aos amigos Felipe Neves que parece pressentir quando preciso da sua amizade, ao Paulo Ricardo meu amigo que acompanhou o início de tudo e que pretendo ser colega, ao Caverna, Zé, Marcelo, Daniel, Fábio e Cassiano. Vocês são meu alto astral, sempre.

Thanks Nick Whiltsher, without your help and friendship I would have certainly not have gone that far.

Thanks to Hernando, Troy, Alex, Sophie, Claire, Camille, Caitlin, Dany, K, Yeri, Lydia, Christina and all my friends from Hyde Park. You all were fundamental during my stay in Chicago. You're my family. I love you all.

Thanks to David Finkelstein, for helping me when I was lost. Thanks to James Conant for accepting me as his student at The University of Chicago. Thanks to Matthias Haase for his advice and philosophical accuracy.

Agradeço à Fulbright e à CAPES pelo incentivo e pelo financiamento desta pesquisa.

Introdução

O Trabalho em filosofia é [...] na verdade um trabalho sobre si mesmo

Na auto-concepção de cada um de nós.

No modo como vemos as coisas

Wittgenstein, Philosophical Occasions PO.161-3

Esta tese nasceu de um gosto pessoal que encontrou um caminho compartilhável com outros filósofos. Reconhecer isso é reconhecer que o mundo acadêmico da filosofia não é, e talvez não deva ser, uma comunidade onde somente aquilo que é a primeira vista objetivamente e intersubjetivamente relevante deve ser estudado. Aprendi que na filosofia as engrenagens pessoais podem acomodar-se aos imperativos das demandas impessoais, pois como acadêmicos de filosofia, mesmo quando há entretenimento e lazer envolvidos, produzimos conhecimento para nossa comunidade conhecer a si mesma e o mundo que a rodeia. Aprendi também que as demandas impessoais ganham vida quando alguém genuinamente se interessa por elas no nível pessoal; quando o que a comunidade precisa apaixonar o homem na sua singularidade.

Esta tese trata de um problema filosófico antigo, o problema cético das outras mentes. Entrei em contato com tal problema filosófico quando fiz o curso do professor Jônadas, em 2013/1, sobre as *Investigações Filosóficas (IF)* de Wittgenstein e o livro *The Claim of Reason* de Stanley Cavell. O problema cético sobre outras mentes, que versa inicialmente sobre a possibilidade de se ter certeza de que outros corpos são habitados por mentes, me chamava a atenção por parecer suscitar um misto de escândalo filosófico e algo essencial a nós humanos. O escândalo era (e ainda é) sentido ao perguntar a mim mesmo como é possível perder esse tipo de conhecimento? Como é possível demandar certeza sobre isso? Já o algo essencial a nós humanos, que ainda sinto presente ao investigar tal problema, parecia vir do exercício de resolver o problema. A busca pela resposta ao ceticismo sobre outras mentes me colocava pouco a pouco em contato com um tipo de resposta filosófica que ao meu ver reencontra as coisas mais essenciais sobre nossa humanidade. Esse algo era o que eu aguardava ansiosamente por descobrir mais e mais.

O curso desenvolveu-se de maneira incrível. Os alunos "devoravam" os textos, buscavam dar o seu melhor nos seminários e creio que poucos foram os momentos que a empolgação filosófica cessou. Tudo isso embelezava o aprendizado. Meu tópico era precisamente sobre o problema de outras mentes, mas ao contrário do que eu esperava, não foi fácil responder a questão que motivava o seminário: *como Wittgenstein compreende a noção de critério? Essa noção permite que ele evite tanto dualismo como behaviorismo? Como*

sua explicações sobre proferimentos psicológicos sobre primeira pessoa e terceira pessoa se conectam?

O texto de Wittgenstein parecia não atacar o problema diretamente, mas o modo como ele aparecia, ao invés de fechar-me em um entendimento dele parecia a obra de um artista que vai pintando lentamente as diferentes camadas de tinta do seu quadro. Aliás, a metáfora de que Wittgenstein pinta um quadro é apropriada não somente pelos estratos de tinta que vão dando origem a um quadro, mas porque Wittgenstein vai oferecendo sobretudo imagens parciais que vão compondo formas de tentação filosófica que ele busca dissolver. É natural que, ao sentir que estamos chegando em uma resposta ao problema, outras alternativas vão sendo excluídas e os caminhos vão se fechando. No entanto, com Wittgenstein, os exemplos de contextos céticos e anti-céticos se proliferavam e era difícil entender o que ele queria dizer com todas aquelas situações de descrição de dor de outros e de proferimentos sobre minha dor e meu corpo.

Wittgenstein explorava a gramática do eu em pensamentos onde o pronome pessoal se conecta à descrição do corpo e de estados mentais. Fazer essa conexão e perceber como os contextos articulavam os diferentes sentidos em que falamos de nós mesmos era o exercício proposto. As maneiras de declarar dor, de expressar dor, de entender a dor do outro apontavam para o modo como aplicamos esses conceitos. Entender a aplicação dos conceitos significava compreender a gramática (termos que veio substituir a lógica, do *Tractatus*) do eu, da dor e do corpo; significava entender o que permitia conectar uma pessoa a outra, isto é, conectar uma mente a outra de modo a vencer o ceticismo.

O primeiro golpe duro onde não deixei de sentir certo gozo foi a frustração de ver que nas Investigações Filosóficas (IF) não havia uma refutação do cético. Pelo contrário, Cavell mais adiante viria com a ideia de *verdade do ceticismo*. Com isso, ele buscava nos deparar com o contexto de enunciação das verdade céticas; um contexto retirado de qualquer situação ordinária, mas ainda um contexto possível para nós humanos. É possível pensar junto com o cético e ver as coisas do seu ponto de vista sem precisar meramente buscar refutá-lo e acusá-lo de proferir frases sem sentido. O ceticismo é uma possibilidade humana e, como nos diz Conant na introdução de seu livro *Varieties of Skepticism - Essays after Kant, Wittgenstein and Cavell*, "o ceticismo ameaça não somente um conjunto de comprometimentos teóricos, mas também - e fundamentalmente - nosso sentido de self, mundo e outro." (p.1)

A "tese" da verdade do ceticismo mostrou-se convincente para mim, mas o antídoto, ainda que seja um remédio de uso contínuo, estava na própria leitura que Cavell faz das IF. Cavell vê em Wittgenstein um filósofo que compreende as raízes do ceticismo. Tal postura filosófica nasce do ceder a tentações intelectuais que, por sua vez, provém de certa imagem que temos do que é ter conhecimento de outras mentes. Essas imagens nascem de confusões

que a linguagem cria quando usada sem respeitar elementos de contexto, ou de sintaxe, ou de semântica, etc. Bem como Wittgenstein dizia em suas investigações preparatórias, no *Blue Book*, as perguntas "O que é significado?" ou "O que é mente?" nos dão a sensação de não saber mais o significado de palavras que usamos cotidianamente, pois o tipo de resposta que buscamos dar visa descrever uma certa essência, que dificilmente dá conta de todos os usos que damos aos termos. Ao perceber que o significado dos termos escorrega às nossas mãos embarcamos em tentativas de respostas que no fundo são motivadas por tentações intelectuais que nos afastam do que é nossa linguagem ordinária no trato com outras pessoas. É na vida ordinária que a linguagem ordinária se expressa e é lá, na maior parte do tempo, que as relações intersubjetivas precisam de pequenos ajustes, sem ameaçar com tanta radicalidade a nossa segurança sobre a existência ou não de outras mentes.

Assim, o diagnóstico que Cavell faz do ceticismo wittgensteiniano é que tal tentação nasce do não reconhecimento do outro -- da perda do outro em sua cotidianidade. O outro que diz na sua frente sentir dor, que grita de dor, mas que diante de uma pessoa cega e presa em uma *puzzle* intelectual só vê um corpo que se contorce desconectado de sensações reais. Esse cético cruel, por assim dizer, abandonou os critérios que nossa forma de vida nutre em práticas ordinárias, o que o leva ele a aplicar conceitos de modo enviesado, a tratar o corpo como "mero movimento", a tratar a mente como "um interno impenetrável", a perder contato com a expressividade da mente pelo corpo.

O conceito de reconhecimento é a solução para não cair na tentação filosófica/filosofante. O ceticismo sobre outras mentes pode brotar nas nossas vidas porque somos capazes de *deixar de reconhecer o outro*. Assim, se o conceito de reconhecimento pode jogar um papel tão relevante na estruturação de um problema filosófico fundamental, me parecia igualmente fundamental estudar tal conceito. Isso é o que se segue ao longo da tese.

Ao longo da trajetória da tese, pouco a pouco, os argumentos irão delinear alguns problemas básicos que desejo ter respondido. A pergunta fundamental que me guia é "O que é reconhecer o outro?", no entanto, para dar uma forma mais específica a esse questionamento, quatro questões serão objeto de exame.

A primeira questão é sobre a unidade de fenômeno do reconhecimento. Tal questão se põe dada a dificuldade que é saber se ao entrar nessas relações existe alguma forma de delinear claros limites para sua ocorrência. O encontro de duas mentes e duas vontades parece requerer um tipo específico de conexão entre pessoas. Se assim for, quais componentes dessa conexão? Conseguimos delinear alguns sentidos básicos em que o conceito de reconhecimento pode ser instanciado?

A segunda questão é subsidiária da primeira. Gostaria de investigar o fenômeno do reconhecimento mútuo. Parece-me que a mutualidade provoca uma forma específica de pensar

o outro. Investigar a forma lógica do pensamento mútuo é algo que acredito ser fundamental para entender a unidade do fenômeno.

A terceira questão é sobre o reconhecimento verdadeiro do que é alteridade. Uma vez que há formas enganosas de reconhecimento onde não há mais que projeção da minha cegueira e das minhas vontades sobre os outros, há claramente um desafio sobre o que é reconhecer o outro verdadeiramente, isto é, na sua diferença.

A quarta questão trata de como o conceito de reconhecimento de outras mentes traz implicações para o problema do reconhecimento de outras vontades. Uma vez que o reconhecimento de outras mentes coloca o sujeito diante de outro ser com vontade, em que medida e como uma vontade deve moldar-se a outra?

Na busca de responder tais questões dividi a tese em duas partes. A primeira parte da tese é dedicada à análise dos argumentos que recorrentemente animaram uma certa tradição filosófica que parte de Descartes, passa por Stuart Mill e chega a Bertrand Russel. Todos eles buscam dar uma resposta ao desafio cético sobre outras mentes. Todavia, seus argumentos irão demonstrar-se insuficientes para afastar totalmente a ameaça cética. Tal insuficiência irá desaguar na necessidade de uma nova análise do problema -- o que nos levará a ver o ceticismo como uma atitude intelectual onde pesa o não reconhecimento do outro.

O primeiro capítulo mostra como o modo tradicional de solucionar o problema cético leva a uma radicalização da questão cética em que a própria capacidade de fazer sentido de si mesmo enquanto detentor de uma vida mental é ameaçada. Início com uma apresentação do problema onde se afirma ser possível descobrir a existência de outras mentes por analogia. As respostas por analogia partem do fato de que posso pressupor em outros a conexão entre mente e corpo, porque observo essa conexão em mim mesmo. Argumentei que o problema dessa pressuposição está em sua fraqueza -- trata-se de uma generalização a partir de um único exemplo. Além disso, trata-se de uma forma inferencial e intelectualizada de conceber a existência de outras mentes, uma vez que no nosso trato ordinário com outras pessoas temos uma experiência mais imediata dos estados mentais de outros. Assim, provas por analogia parecem desfazer certa imediatez que experimentamos no trato com outras pessoas.

No entanto, se fossemos ficar com o fracasso das provas por analogia em mãos, a questão que salta aos olhos seria se faz sentido pensar que minha mente é a única que pode existir. Em caso positivo, faz sentido distinguir minha mente do mundo e dos outros? Para responder tal pergunta tingida de solipsismo, mobilizei o argumento da linguagem privada e da possibilidade de seguir regras presente em *Wittgenstein - On rule following and Private language*, de Kripke. Tal livro foi um ponto de virada nos estudos sobre Wittgenstein, mas também nos estudos sobre o ceticismo. O que o cético de Kripke desejava era um fato que

pudesse mostrar que nossas práticas de seguir regras são fundamentadas na realidade. Tal fato, no entanto, não é jamais encontrado por Kripke e ele conclui que o paradoxo só aceita uma solução cética. Em sua solução cética Kripke diz que o máximo que podemos obter é justificação para nossos proferimentos baseada na constatação de que linguagem é uma prática sustentada pela comunidade, de acordo com critérios de aplicação de conceitos e visando expressar as necessidades da nossa forma de vida. Nessa imagem Kripkiana, a linguagem não dá a conhecer a vida mental do outro, mas é uma mera moeda de transações mentais. Esses três componentes, acordos, formas de vida e critérios, são os principais pilares sobre os quais é possível afirmar que a linguagem é um bem público. Assim, enquanto um bem público os conceitos de mente, corpo e pessoa são a base para fazer sentido do mundo que nos rodeia, dos outros e de mim mesmo. A publicidade da linguagem mostra que para entreter pensamentos em solidão, retirados de qualquer diálogo, é necessário um recurso a uma noção de alteridade básica com a qual os laços de linguagem são criados.

Se por um lado Kripke não oferece um fato que satisfaça o cético, por outro Crispin Wright acredita ter esse fato em mãos. Ele defende que constituímos o sentido de nossos proferimentos em atos de assentimento retroativos. Para ele, o sentido é constituído por um sujeito cuja autoridade é garantida por condições sabidamente propícias em que ele assente o conteúdo inarticulado e bruto das suas ações e pensamentos.

Ainda que essa solução dada por Wright tenha a autoria como aspecto importante daquilo que reconhecemos como sendo fundamental na prática da linguagem, essa solução será criticada por John McDowell. Este nos oferece uma crítica ao constitutivismo de Wright, uma vez que ela recai nas dificuldades que concepções aditivas do significado implicam. McDowell propõe que adotemos uma concepção disjuntivista de significado, onde intenções e significado andam juntos, sem atos recursivos de assentimento. Ele defenderá que essa unidade é um critério formal da atividade cognitiva humana. O importante dessa concepção sobre o significado é que ela se conecta com o problema conceitual de outras mentes. A possibilidade de conceber outras mentes é parte inexorável do processo que é conceber a mim mesmo e o tipo de ser que sou.

No segundo capítulo busquei explorar diferentes teorias acerca de como se dá a aquisição do conceito de outro. O primeiro caso é o da chamada “Teoria da Teoria”, que afirma que temos uma proto-teoria acerca de nossa vida psicológica. De acordo Paul Churchland, as regularidades comportamentais que observamos em outros e em nós mesmos permitem que criemos certas noções psicológicas comuns a nós e a outros através das quais nós tacitamente conhecemos a nós e a outras mentes.

Os problemas com essa concepção chamada Teoria da Teoria nos levarão a considerar a Teoria da Interação. Esta, por sua vez, é uma teoria que tem suas raízes nos

estudos de psicologia de J. J. Gibson e consideram os avanços da neurociência essencial para compreensão o fenômeno da intersubjetividade. O essencial a teoria da interação é que ela afirma ser possível perceber diretamente emoções e fenômenos psicológicos em outros através de uma noção mais robusta de comportamento -- a ideia de comportamento expressivo que já está presente em Wittgenstein, mas que é retomada de modo a dar mais espessura. De tal modo, a ideia de reconhecimento depende de que o sujeito possa constituir os mecanismos de percepção direta de estados mentais de outros, constituí-los significa ser capaz de gerar uma fruição no trato social.

Posteriormente passo à análise dos argumentos de Thomas Nagel e sua leitura de Sartre. Nagel acredita que a teoria sartreana tem a virtude de mostrar que nosso conceito do outro advém de um momento primordial no qual nos sentimos objeto diante do olhar do outro. O olhar do outro será o que Sartre chamou de "cogito do outro" uma espécie de teoria onde o outro é joga um papel no nosso mundo psicológico. O conceito do outro nesse caso é estruturado primariamente pela consciência de ser objeto do olhar do outro.

Finalmente, as ideias analisadas no fechamento do segundo capítulo farão a transição entre o que é uma investigação de resposta ao cético e uma investigação sobre o que é o conceito de reconhecimento. Nesse momento final busco mostrar a diferença que há entre *conhecer* e *reconhecer* outras mentes. A análise de Stanley Cavell em "*Knowing and Acknowledging*" mostra como o problema cético sobre outras mentes emerge a partir da não observação da gramática dos enunciados psicológicos em primeira pessoa e terceira pessoa. Os modos como falamos dos nossos estados psicológicos e o modo como falamos dos estados psicológicos de terceiros guardam uma desanalogia que confunde aquele que acha que se pode encontrar um forma de conhecimento comum a ambas. Elas inspiram modelos do que é conhecer uma mente, mas ambos modelos são incongruentes e por isso a exigência cética se torna insuperável. Sem adotar nenhuma das possíveis respostas, Cavell irá propor que a relação de reconhecimento é o modo que assumimos no trato intersubjetivo. Essa relação é vivida sob a forma de endereçamento (*address*) direto à pessoa e ela será o objeto da segunda parte desta tese.

A segunda parte da tese, composta pelos capítulos 3 e 4, é onde irei explorar mais detalhadamente o conceito de reconhecimento. Eu inicio expondo dois autores que representam duas formas contemporâneas do pensamento hegeliano acerca do conceito de reconhecimento. Tal conceito mostra como a estruturação de relações de reconhecimento é dependente de dois atos: o primeiro é a validação social de um ser humano através do olhar. Validar alguém socialmente é uma forma de percepção avaliativa -- um modo específico de perceber o que é um ser humano. O contraste com o que é perceber algo não humano dá origem ao conceito de triangulação básica. A triangulação básica ocorre entre dois sujeitos que

alternam posições passivas e ativas de poder e objetos inanimados que entram nas trocas subjetivas como instrumentos de expressão de racionalidade prática. Essas ideias demandaram a exploração de algo básico nas relações de reconhecimento: as propriedades de endereçamento e direcionalidade, uma vez que o olhar é já uma primeira forma de manifestação do que é dirigir-se a outro ser humano.

Se reconhecer o outro é um ato de validação social, então Thomas Reid e sua teoria dos atos sociais da mente é certamente um filósofo no qual poderemos encontrar critérios para delinear e aprofundar o estudo do que vem a ser o ato de reconhecer o outro. Reid falará de três condições fundamentais que devem ser atendidas: a participação de dois polos em trocas subjetivas, a expressão de um ao outro e o reconhecimento dessa expressão por parte de um e outro.

Tais condições serão desafiadas na literatura contemporânea sobre o que está envolvido em relações de reconhecimento. Christopher Peacocke mostrará que a lógica envolvida nos atos de reconhecimento é dependente da articulação do pronome pessoal de primeira pessoa e do uso de demonstrativos. O pronome pessoal de segunda pessoa, nas suas palavras, não passa de um componente da linguagem comum analisável nestes dois mais mais básicos. A dificuldade, no entanto, é a imagem do que são atos de reconhecimento segundo a proposta de Peacocke. Para ele, as relações intersubjetivas se estruturam através de pilhas de pensamentos que parecem ser vividos individualmente. Falta a Peacocke uma certa consideração de que o jogo do reconhecimento é jogado com a presença corporificada dos agentes é necessária.

Uma crítica ao modo como Peacocke concebe relações de reconhecimento virá de Sebastian Rödl, que buscará mostrar que no fenômeno de ação conjunta (*joint action*) não é necessário recurso a qualquer tipo de pensamento demonstrativo. Para ele, o reconhecimento permite pensar a primeira pessoa do plural onde o outro não é tomado meramente como "aquele que me escuta", como se mero receptor, mas alternativamente coloca ambos agentes em endereçamento mútuo.

Finalmente, uma última forma de compreender a propriedade de endereçamento é encontrada em *On Address*, de Adrian Haddock. Haddock nos oferecerá uma imagem diferente do que o endereçamento não-intencional. Sua ideia é pensar o corpo humano como uma espécie de ato de endereçamento, mesmo quando não intencionalmente engajado em trocas intencionais com outros. Desse modo, o corpo é um emblema da vida humana, do que fazemos de nós mesmos.

O tema do endereçamento direcionado a outros, ao final do terceiro capítulo assume a como parte de sua problemática a capacidade de endereçar a si mesmo. No capítulo final da

tese, irei buscar mostrar como as teorias éticas contemporâneas do reconhecimento tem concepções de endereçamento jogando um papel central nos seus argumentos.

O último capítulo da tese explora a ética da relação de reconhecimento. A intenção principal é mostrar que além de perceber o outro como dotado de uma mente e de inteligência, o outro também é visto como um ser desejante. A trajetória do capítulo mostra que é pela nossa capacidade de endereçar-se a outros que a forma da moralidade é conhecida por seres racionais como nós.

Eu inicio o capítulo final apresentando as considerações de Nagel em “War and Massacre”. Esse artigo elabora uma interessante discussão dos princípios consequencialista e deontologista sobre a moralidade. Nagel elabora contextos em que tais princípios não conseguem reger por si só a forma do agir, pois a capacidade de agir conforme o dever ou a capacidade de agir conforme a máxima utilidade não podem ser atualizadas sem inúmeras concessões. Assim, diante de um cenário onde os princípios não podem atuar na pureza de suas prerrogativas, Nagel aponta para a prática de dar razões e justificar as ações. Essa prática na vida moral humana é regida pela capacidade de endereçamento através da fala e representa em si mesma uma forma de respeito à dignidade do outro, mesmo quando a justificativa é para alertar sobre os erros morais que alguém possa cometer.

Apoiado em tal premissa, ofereço os argumentos de Stephen Darwall apresentados em *The Second Person Standpoint*. Nesse livro Darwall busca mostrar que a forma da moralidade pode ser derivada da forma do endereçamento. Ele busca na teoria dos atos de fala de J. L. Austin e na teoria das atitudes reativas de P. F. Strawson as condições para derivar o imperativo categórico (IC). Ele acredita que assim como os atos de fala, os atos de endereçamento tem condições de felicidade, isto é, condições materiais e de legitimidade que conferem "existência" a tal ato. Já Strawson é chamado a corroborar a ideia de responsabilidade moral, uma vez que as atitudes reativas são reflexos de algo que está na base delas: a capacidade do sujeito de se pensar como moralmente responsável. O imperativo categórico, nesse caso, viria da forma do endereçamento, pois ao endereçar-se a mera presença hipotética de um humano no outro polo da relação de reconhecimento imprime os requisitos formais de como uma pessoa deve agir. Darwall mostrará que Kant chega a formular na *Crítica da Razão Prática* uma certa aproximação com a ideia da forma do Imperativo Categórico ser determinada pela forma do reconhecimento da humanidade do outro.

A última autora a ser apresentada na tese é Catherine Korsgaard. Seu livro *Self-Constitution - Agency, Identity and Integrity*, exhibe teses igualmente sobre como o conceito de reconhecimento joga um papel fundamental nas nossas práticas morais. Diferentemente de Darwall, Korsgaard acredita que a forma do endereçamento não é dada por um outro hipotético com o qual pensamos o Imperativo Categórico. Para ela a forma da moralidade é dada pela

auto-consciência, pois é através dela que endereçamos a nós mesmo sobre nossas ações, identidades pessoais e integridade moral.

PART I
Chapters 1 and 2

Chapter 1 - The Problem of Other Minds

"I believe that he is suffering." — Do I also believe that he isn't an automaton?

"I believe that he is not an automaton", just like that, so far makes no sense.

Wittgenstein, Philosophical Investigations

1. From "Me" to "Others"

Self-consciousness is immediately manifest to the kind of beings we are. We know we are self-conscious by simply experiencing the awareness of being conscious of something. However, among the things we are conscious of there are beings that look like us, other humans that talk about their experiences, thoughts, feelings and life. We immediately think them as also self-conscious, we tend to not put into question the fact that others are also beings that experience consciousness like us.

The experience of sharing routes of interest, sense of humor, narratives, what to love, what to hate, the senses of beauty, of what is immoral, the sense of outrageousness, etc, is what gathers humans into groups. Nonetheless, it's not always that communication flows unproblematically. Sometimes we disagree with others in ways of feeling, thinking and experiencing the world. As a consequence we feel alone and our social bonds seem to be the work of a fragile miracle. When the bonds among humans are broken hatred, loathing and distaste for others can blind us to the other's humanity. The experience of this distance or closeness to other humans is at the basis of what modern philosophers have called skepticism about other minds.

The aim of this dissertation is to examine the skeptical problem of other minds. There are many formulations to this philosophical problem: It goes from the radical doubt about the existence of other minds to the less radical difficulty in knowing how other cultures, genders or identities experience (perhaps only parts, perhaps the whole of) a shared world. My strategy addresses the radical problem, since it contains what is more general and common to all other variations.

In this chapter we will understand how the epistemological problem of other minds leads us to the conceptual problem of other minds. First we will examine why analogical arguments

fail to give us an answer about the existence of others and, from this incapacity we will derive a more substantial picture of what is involved in the problem -- the impossibility of a private language as a individual and inner linguistic phenomenon. The investigation of the possibility of private language is a facet of the rule-following problem. This problem will confront us to the difficulty of teaching someone to obey and learn how to follow rules. In this scenario, we will understand better the nature of language and what is to communicate with other minds. If meaning were a private entity "inside one's head", then this would entail the privacy of the experience of meaning; and this is all the skeptic needs to claim the truth of skepticism. If meaning is not private to one's mind, but a public construct, then all ordinary concepts, the concept of mind included, are crafted in public and shared activity. Surprisingly, our analysis of the private language will lead us to a different idea of what language and communication are, since the threat of skepticism will be understood in its motivations and procedures.

2. The (not so) old Problem of Other Minds

Skepticism is one of the oldest philosophical postures. Ancient Greece was the birthplace of both Academic skepticism and Pyrrhonism. Academic skepticism developed within Plato's Academy under the teachings of Achesilaus and then it was revitalised by the dialectician Carneades. This school of skeptical thought was a reaction to the Stoic dogmatism, who believed that, to attain happiness, one should act on the basis of knowledge. Academic skeptics also believed that the goal of life is to conduct one's life so as to attain happiness, nonetheless, they believed that there's no criterion for determining truth. Truth is impossible to achieve since we are liable to illusion, deception and other forms of mistake. Hence, no knowledge is possible and no true happiness should depend on knowledge. Thereby, as a reaction to Stoic's dogmatism, we should guide ourselves following what is merely plausible suspending judgment about what is true.

Pyrrhonism is a more radical form of skepticism, for Pyrrhonists don't make concession to the plausible or reasonable. They simply follow the consequence of not having beliefs. Their aim was to purify their souls by suspending judgment (epoche) in order to attain tranquility (ataraxia). After suspending belief and judgment about reality, the Pyrrhonean can live his life and act according to the tranquility achieved.

Anita Avramides, in her book *Other Minds*, claims that there were some skeptical concerns about other minds in Ancient Greece. Aenesidemus attributes to Agrippa the study of the modes of giving rise to beliefs. In this study different modes of belief formation are studied so as to consider the conditions under which the agent form his judgments. The classic example is the case of perception. The taste of honey can be felt as sweet or bitter depending on the

health of a given person. Also rectangular or circular towers can be perceived differently according to the point of view. The problem for Greek Skepticism seems to be why or how can we experience differently the same world. In contrast to the modern skepticism about the existence of an external world, Burnyeat shows that ancient skepticism

[...] leave untouched — indeed they rely upon — the notion that we are deceived or ignorant about *something*. There is a reality of some sort confronting us; we are in touch with something, even if this something, reality, is not at all what we think it to be. Greek philosophy does not know the problem of proving in a general way the existence of an external world. (Burnyeat, apud Avramides, 2000, p.27)

This feature of ancient skepticism demonstrates that Ancient philosophy acknowledged that different persons have different perspective and experiences on the same things and that was taken as the main reason for skepticism about the possibility of knowing the truth. Even though this kind of skepticism demands a certain articulation between mind and world, this wasn't enough to motivate skepticism about the existence of an external world.

Naturally the concept of mind was also a also slightly touched as part of the problem of knowledge. Sextus reports the Cyrenaics as making the observation that “each person is aware of his own private *pathos*, but whether this pathos occurs in him and his neighbour from a white object neither can he himself tell, since he is not submitting to the pathos of other person.” (idem. p.28) This is sometimes sufficient to motivate skepticism about other minds in a modern conception, but here, as in the case of Academic skepticism and Pyrrhonism, there's no question about the existence of the other mind. The doubt here is simply whether people experience the same thing privately, but there's no question about the existence of a mind that experiences.¹

The question of the existence of minds and world is a modern one. Ancient skepticism is different from modern skepticism. Modern skepticism is more radical than Ancient skepticism and expresses a different human possibility. The problem of the existence of a world and of other minds is usually recognised as showing up for the very first time in Descartes *Meditations*. After stating his famous argument about the undeniable self-evident existence of some subject of thought as a result of his method of hyperbolic doubt, the “*cogito, ergo sum*”, he runs to a supplementary difficulty that necessarily remains open after his conclusion: the impossibility of knowing the existence of other minds with the same level of certainty that I know my own

¹ Avramides quotes Voula Tsouna “ancient denials that we can have access to the path [experiences] or thoughts of our neighbours are weaker than similar claims occurring in modern discussions about other minds” (Tsouna in Avramides 2000, p. 29)

existence as a thinker. Few pages later, after the conquering back the knowledge about the existence of an external world, he addresses the problem of not knowing whether others are thinkers as he is:

If I look out of the window and see men crossing the square, I normally say I see the men themselves [...] Yet do I see anything more than hats and coats which could conceal automatons? I judge that they are men.
(Descartes, *Meditation II*, AT 32)

The Cartesian scenario is a rich source to understand the problem of other minds, even though he does not formulate it with the precision we see later in the history of philosophy. The way he describes other humans from his point of view is such that facilitates the adoption of a skeptical posture. Looking at people from the window, from afar, in such a way that we can't see their expressive faces and bodily particularities puts the observer in a distanced relation to them. In such a position no second person relations are possible, for no interaction is actually created. The reader of the *Meditations* is guided to imagine herself observing other humans from a distant point of view that tends to objectify an otherwise natural human intercourse. Descartes conduces us to a subtle shift in the way we look at others. In such a position I can only *judge* that others are humans, *infer* they are people like me.

The call for a judgment gives us precisely the indication that the problem of other minds has been given an epistemological guise. To know about the existence of other minds it is necessary to *think a true proposition* about that mind, and to be able to do that it is necessary to understand *which premises* and *on what kind of evidence* I can rely to consistently state the existence of others. The problem of knowing the truth about the existence of other minds and determining upon what kind of evidence I came to know about those very truths is what I will call "the epistemological problem of other minds".

The epistemological problem presupposes the intelligibility of the idea of others having minds -- for whatever proposition stating that others have minds must be first and foremost *understood*. So the concept of other minds must make sense even before I can ask the epistemological question. The problem of *understanding* the sources and the meaning of the concept of other minds is what I will call "the conceptual problem of other minds".

We will start now by examining the answers to the epistemological problem of other minds, but, I hope it will get clear as we progress, this chapter's aim is to connect the epistemological problem to the conceptual problem of other minds.

3. Arguments from Analogy

There are two paradigmatic ways of answering to the problem of other minds. The first is called the *negative thesis* and the second is the *positive thesis*. The negative thesis consists in having as a premise one's own case and then move to an account of the other based on our knowledge from first personal perspective. The positive thesis, on the other hand, consists in taking the knowledge of others as basic as my knowledge of myself.

The negative thesis, commonly represented in analogical arguments, is the oldest strategy. It prevails during the modern times, but only in the 20th century has been rejected as a bad strategy. The first clear formulation of this strategy was given by John Stuart Mill. He says:

I conclude that other human beings have feelings like me, because, first, they have bodies like me, which I know in my own case to be the antecedent condition of feelings; and because, secondly, they exhibit the acts, and other outward signs, which in my own case I know by experience to be caused by feelings. [...] We know the existence of other beings by generalization from the knowledge of our own (Mill, J.S.,1889 243)

Mill's argument gives us the general structure of arguments from analogy. It starts stating our shared bodily nature. Then it recognizes one's own body as the substance that is the condition for a bearer of feelings. Then, it takes the knowledge from one's own case as the basis for a generalization to all bodily resembling creatures.

Mill's argument was an answer to Reid's characterisation of the problem of other minds as disconnected to epistemology. For reasons that are not important now, Reid believed we cannot rationally have a general conception of mind, but Mill disagreed. Mill tried to show that with an analogical argument it is possible to conceive others like me. Hence, the concept of mind is universal. Unfortunately, as we will see later, his argument's mistake resides in the weak generalization.

3.1. Russell's Knowledge by Description

Mill's spirit was still present in the 20th century philosopher Bertrand Russell. In his 1905 article *On Denoting*, Russell argued that we can obtain knowledge about other minds by means of denoting phrases. There are many propositions whose truth we don't know by acquaintance with the fact that enables verification, but we can only know by description. Descriptions are phrases that denote objects, and this is a mechanism of our language that allows us to refer to things that are not immediately presented to us. In Russell's words "The distinction between *acquaintance* and *knowledge about* is the distinction between things we have presentations of,

and the things we only reach by means of denoting phrases" (RUSSELL, 1905, 479). So, objects like the present king of France or the earth's center of gravity can be referred to by means of definite descriptions, since we cannot have acquaintance with them, while my own thoughts, feelings, wishes and sense data from objects of perception are known by acquaintance.

Knowledge by acquaintance and Knowledge by description will be characterised as two forms of experience: mediate and immediate. In *The Problems of Philosophy* (1912), Russell says that acquaintance is a sort of *immediate* knowledge. If on the one side, our thoughts, feelings and desires are known by awareness of our inner states, on the flip side, he believes that our knowledge of external objects -- such as chairs, the planets and human bodies -- is a sort of *mediate* knowledge. (In Russell's view I have direct awareness of shades of color, the so-called sense-data, it is from such brute information that I make up the appearance of the object to my consciousness.) In the same vein, his prior view on the problem of other minds is that "There seems to be no reason to believe that we have ever acquaintance with other people's minds, seeing that these are not directly perceived: hence what we know about them is obtained through denoting." (1905, 480).

Therefore, Russell believed that our knowledge of other minds is only possible by means of denotation. To denote singular objects that we are not immediately presented to, we have make use of definite descriptions. Russell says definite descriptions occur in sentences where a free variable and an existential quantifier are used to refer to one object without naming it explicitly, but making use of descriptions and judgments of identity so as to isolate the object and define it. So when we say "The actual king of France is bald" the analysis of this sentence would present underneath its surface grammar quantifiers and variables that claim the existence of *some* object, and moreover *one and only* object, that has such and such properties. The case with other minds would be analogous, for a definite description would allow us to denote other mind's content such as feelings, thoughts and sensations.

When I desire food, I may be aware of my desire for food; thus "my desiring food" is an object with which I am acquainted. Similarly we may be aware of our feeling pleasure or pain, and generally of the events which happen in our minds. This kind of acquaintance, which may be called self-consciousness, is the source of all our knowledge of mental things. It is obvious that it is only what goes on in our own minds that can be thus known immediately. What goes on in the minds of others is known to us through our perception of their bodies, that is, the sense-data in us which are associated with their bodies. *But for our acquaintance with the contents of our own minds, we should be unable to imagine the minds of*

others, and therefore we could never arrive at the knowledge that they have minds. (Russell. 1912, p.37)

Now the last step is to relate by means of an analogy acquaintances and descriptions: the immediate knowledge I have of my inner life, and the immediate knowledge I have of the other's body, with the mediate knowledge I have of my body with the mediate knowledge I have of the other's mind. In other words, Russell believed the connection between my sensations and my bodily behaviour is analogous to the relation between the body of others and their minds.

3.2. Two other attempts

After Russell, many different philosophers tried to settle the issue by means of analogical arguments. Malcom in his famous article *Knowledge of Other Minds* mentions two arguments that we cannot fail to cite. The first, claimed by Stuart Hampshire, inverted the poles of the classic analogical argument. Instead of starting from the first person, he believed that starting from the second person we could warrant the validity of the inference. So, roughly speaking, when someone infers that I'm dizzy by analysing my behaviour, this person will react treating me as if I were dizzy: she will ask me whether I'm well, and will ask me if I want to throw up, etc. If I end up at the hospital and get better, I'll know that dizzy is the concept for what I was feeling. Thus, when the inference about my mental states is good, I can rely on it to learn by myself the good inference-procedures. Next time, when I see someone displaying dizzy-behaviour, I'll know what to do, which kinds of mental states are going on in the case of this person. So, Hampshire thinks I can use myself in a reliable way to draw conclusions about others, because I have learnt how to think others' minds comparing their inferences to my own inner states. The idea is: I first am subjected by others, they name my inner states and they represent me to myself. Then I can analogously know what happens in other's minds based on my former experience of subjection. The second argument comes from Price. He says in *Our Evidence for the Existence of Other Minds* that "one's evidence for the existence of other minds is derived primarily from the understanding of language" (p.429). He says that we all know from introspection that when we speak the sounds we make are "symbols in acts of spontaneous thinking"; therefore similar combinations of sounds spoken by others "probably function as instruments to an act of spontaneous thinking, which in this case is not my own." (p.446) So from the fact that others speak meaningful language, we can deduce that they are intelligent creatures.

Malcom considers both arguments cases of analogy, even though they invert the classic starting pole of the relation. I agree with his analysis that inverting the pole does not prevent the argument from the needs of a first person standpoint, but just makes it appear later. Supposing

that the first person perspective is always the perspective of the ignorant about the existence of other minds, we see that all these arguments will in some moment lead the reader to adopt the position of ignorance. So, on the one hand, the case of Hampshire, where the inverting poles seemed to free us from the dependence of the first person premise, disguises the actual need for an inference from the perspective of the first person. Since in his example I learn to make inferences by learning how others draw inferences about me, I have to make the reverse process of inferring from my own case to other's. Hence, this argument only delays our need for a first person perspective in the analogy, but does not avoid it. On the other hand, Price's argument also commits the same sin. The supposition that other minds exist because other persons speak an intelligible language is made on the basis that I speak an intelligible language. So, as a speaker and knower of my own experiences I analogously attribute the same nature to others. Nonetheless, this would be an argument from analogy using language, and not bodily behaviour, that would still take the first person bridge to others.

Now, what are the problems with the argument from analogy? It seems that many problems can be adduced for these arguments. First, if we understand the analogy as a case of induction, it seems that this is too weak a form of induction since it rests only on one (my own) specific case. A second problem is that usually inductions allow us to test the correctness of the inference, but if I can never have direct awareness of someone's mind, I can never truly test it. Thirdly, it seems that whatever counts as an evidence for my knowledge of other minds does not necessarily serve as evidence I have for the existence of my own mind, for they can differ not only in number, but also qualitatively. So, when I generalize from my own case I maintain the concept of mind without knowing if it is appropriate for the other's case.

If arguments from analogy have persisted throughout the history of philosophy, how come that people have never noticed its weaknesses? Perhaps there's a hidden reason for all those philosophers to persist in the same strategy and I think Wittgenstein was the one who was able to decode those reasons. Instead of sheer refutations and exploring the argument's weaknesses, Wittgenstein tried another method for philosophy. His method consists in going as far as the argument allows us, to see which kind of ground presupposition it is based on. He explores the logic that structures the philosophical problems so as to see in their roots the kind of answer they demand. This kind of method not only shows why an argument does not work, but it also reveals aspects of our rationality -- temptations, partial perspectives, intellectual vices. As a consequence, the analysis of philosophical problems bring a sort of self-knowledge that reveals a shared intimacy, the intimacy we access because we are rational beings.

In what follows, I want this chapter to show what is the philosophical difficulty that underlies analogical arguments and how this leads us to the conceptual problem of other minds. If I'm allowed to treat them as whole, arguments from analogy inadvertently presuppose a

separation, a gap that must be filled. The gap goes from one's mind to the knowledge of the other, but to understand this it is necessary to understand how meaning one's own mind is ever possible. The thread I want to follow leads from the impossibility of knowing the other by means of analogy to the problem of not knowing anymore what the concept of mind means when I use it to talk about my own experiences.

4. The Conceptual Problem of Other Minds

Analogy is the main form of the negative thesis about other minds. As we have seen, most forms of analogical argument about other minds sooner or later will need to make appeal to the first person perspective, where one assumes one's own experiences as private and certain. From this warranted perspective philosophers draw the inference to obtain knowledge about third person states. Last section shows that the basis for my induction about other minds is weak - among many problems, the most remarkable is the bad generalization from one case, and yet from my own case. The consequence is that I don't know if the concept I have from *my mind* also applies to the *other's mind*. So, if on the one hand analogical inference fails, on the other hand it actually brings about more problems than the ones we started with. Once we notice our concept of mind is weak, we must investigate the conditions on which we can speak of a public and shared concept of mind. For, if there's one public concept of mind, we can at least know what are we talking about.

The idea of a public concept of mind comes from Wittgenstein's (so-called) "Rule Following Argument". This argument is found in Wittgenstein's *Philosophical Investigations* (PI), in the sections that go from §185 to §243 and is basically an investigation about meaning, mind and normativity. Much of the recent debate about the PI has been dedicated to understand what exactly the rule following sections are trying to say, but here I won't focus on this debate. I'll deal instead with a related traditional problem that became important in analytic philosophy: the publicity of meaning.

The rule following argument starts with the description of a teacher trying to help a pupil learn the mathematical operation of addition. The teacher asks his student to add 1 to the series of natural numbers starting from 0. As the the pupil goes on producing the series he fatally arrives at the number 1000. Then, simply to challenge the student, the teacher asks him to add 2 to each number following 1000, so as to obtain 1002, 1004, 1008, etc, but at this point something awkward happens. Instead of obtaining the result 1002 immediately after arriving at 1000, the student's result is 1004. After seeing this, the teacher says:

[...] "Look what you're doing!" -- He doesn't understand. We say, "You should have added *two*: look how you began the series!" -- He answers,

“Yes, isn’t it right? I thought that was how I *had* to do it.” — Or suppose he pointed to the series and said, “But I did go on in the same way”. -- It would now be no use to say, “But can’t you see . . . ?” -- and go over the old explanations and examples for him again. In such a case, we might perhaps say: this person finds it natural, once given our explanations, to understand our order as we would understand the order “Add 2 up to 1000, 4 up to 2000, 6 up to 3000, and so on”. (PI 185)²

How should the teacher proceed in this situation? It seems that the student does not notice his mistake. He is pretty sure of what he is doing. He answers naturally that the result is 1004 and this is precisely the naturalness we were expecting, but obtaining the result 1002. Our problem now is: How can I explain to this person, who seems so convinced, that the result is wrong? How can I tell him that the step he should have taken was a different one? And as I see myself explaining how to follow the addition rule, I am also able to ask myself: what, at any stage, is to be “in accordance with a rule”?

5. Kripke and the Rule-Following Problem³

In *Wittgenstein on Rules and Private Language* (1982), Kripke tells us how the sections of PI strikes him. He believes Wittgenstein introduced a paradox about the possibility of following rules that doesn't have a positive solution, but only what he calls a "skeptical solution". His analysis of the possible answers to the rule following problem leads us to conclude that Wittgenstein has as a skeptical argument against the possibility of knowing how to follow rules. Kripke claims that at each step we take in following a rule, nothing can tell us we did the right move. Therefore, there's no sense in questioning which step can be the right one, for every course of action is consistent with the steps previously taken. We will see that the best rules can do for us is to justify us in meaning what we say, which is different from giving us actual knowledge of the right step to take.

Kripke justifies his claim calling our attention to PI 201, where Wittgenstein announces a paradox about rule-following.

This was our paradox: no course of action could be determined by a rule, because every course of action can be brought into accord with the rule.

² Wittgenstein’s *Philosophical Investigations* will be quoted always according to the model (PI XXX).

³ The Rule Following debate that I present here is pretty much the result of the teachings of David Finkelstein's course "Late Wittgenstein" attended during the fall of 2016.

The answer was: if every course of action can be brought into accord with the rule, then it can also be brought into conflict with it. And so there would be neither accord nor conflict here. (PI 201)

If this paragraph is the epitome of Wittgenstein's discovery about rules, then what we have in our hands is a skeptical conclusion. As Kripke points out, suppose you're asked to imagine a computation you have never done before: "68+57". You would obviously, I expect, come to the result 125. This is hardly a matter of dispute, for you have often used the symbol "+" and have made many computations before. But now Kripke asks you to imagine a bizarre skeptic that comes to you and claims you are misinterpreting your own previous usages. The weird skeptic says the result is actually 5. You ask him "Why?", and after the outrageousness fades away, your primary reaction would be to send him back to school in order to finally learn how to add numbers. But imagine he continues, dead serious, saying "For the case of 57 + 68 precisely, you should answer 5". This seems absurd, for you naturally think you know that the function you applied in the past will give you the result 125. However, the skeptic brings up a real threat when he asks about the authority and the grounds on which his opponent is based. "Who is to say which function was that you were previously using?" -- he might ask. As Kripke notes, in the past I have given myself only a finite number of examples of the function I called addition. What if all steps I have previously taken in following a rule are actually the same steps of a different function called quaddition? It happens that all my previous applications of the rule coincide with the exact experience of quadding those numbers. How can I know which steps I had actually followed? Suppose I was unaware of this resemblance and suppose the skeptic is showing me a function that is totally consistent with my previous steps. He shows me

$$\begin{aligned} x \oplus y &= x + y, \text{ if } x, y < 57 \\ &= 5 \quad \text{otherwise.} \end{aligned}$$

"So the result of 57 quus 68 is 5." - The skeptic says. You, as the skeptic's opponent, know all your past occurrences of the function 'plus' yielded the correct results, but all those results are also consistent with quus. It could have been the case that when it comes to 68 and 57, specifically these two numbers should be quadded, you just didn't know it. Now suppose you insist. You say: if I am adding 68 and 57 the function "plus" says you should obtain 125, but if I am quadding, the 'quus' function says you should obtain 5. But now the skeptic is no longer pressuring you to accept the result as 5, he remains waiting for you to realize that you can't tell the difference in what you were doing before. No matter what, you can either answer 125 or 5,

for both results are consistent with all your past usages -- as Wittgenstein's paradox is supposed to show, both outcomes are consistent with the previous steps, and, in principle, infinite others are too.

Yet, the skeptic questions "Can you say why you were not using the function Quus?" Of course as the opponent, you will insist that you were not, and that there's no misinterpretation about your own past usages; you will argue that you remember always having proceeded according to the addition function. You remember always having added each number after one, and the procedure should guarantee the application in the present case. The skeptic, on the other hand, will only drop his last challenge: "Could you give me a fact that shows I cannot adopt 5 as an answer?"

Kripke's conclusion will be that there are no facts that can bar the agreement of the function quus with all other past uses of "plus". To arrive to this conclusion, he shows that three good answers will fail to dismiss the skeptical threat.


5. 1. Three shots against Skepticism

The first answer to the skeptic is the rule thesis. It consists in the idea that, before challenging me, I had a rule, so I can remember using a rule. Imagine that I answer the skeptic saying "I have a rule that tells me how I should proceed in this case. It is a rule I still remember from when I was a kid. My teacher taught me to add marbles and I still remember that I countless times gave the right answer. So now, all I must do is to get 57 marbles and then add to them 68. That's my rule!" To this the skeptic could answer "Yes, and then you should leave only 5 marbles on the table. For this is the next step -- which by the way is totally consistent with your previous calculations". Surprisingly or not, the skeptic is still holding his belief. But why? The skeptic is proposing that we should also reinterpret all the past rule applications as completely consistent with quus. The rule I claim to have is actually a procedure that could be of any sort. Could be marbles, could be dots in a paper sheet, could be anything. The procedure itself is what the skeptic means to be reinterpreted as possibly meaning quus all this time. So, you can't really know if what you were doing was quus or plus all this time, because everything that you have learnt can be reinterpreted so as to be consistent with your new step.

The second interpretation Kripke faces is Platonism. For a Platonist, meaning is an objective entity, like objects of pure thought that we access only because we are beings capable of thought. This view is seductive because of its strong appeal to the idea of rules as rails that lead to infinity, as Wittgenstein suggests in PI 218. This idea is sort of an image evoked by Wittgenstein to express the feeling we have that following a rule is to do something already determined. But if what I'm doing is determined by a rule, the next step of the same rule is also necessarily determined, and so on. Hence we conclude that rules have the power to predict

infinitely their next steps, as if all the infinite possible steps were already waiting for us. Hence, whenever someone follows a rule what he is doing is simply to instantiate what that rule determine him to do. Therefore, when we grasp the meaning of a rule we are actually accessing a platonic entity.

This view, even though extremely alluring, incurs in some pretty serious problems. First, how do I know whether the entity that you have access to is the same I have access to? Maybe when I say "tomato" the spoken sign will lead you to pick the wrong platonic entity in the realm of meanings, but how can I know that you didn't? Hence, in Platonism, simple acts like conversation become highly suspicious and miraculous. Second, if rules teach us how to interpret written or spoken signs, the Platonic meaning is supposed to lead us to the true meaning of those signs. But, it is shocking to notice that whenever I interpret a sign, I think again by means of new signs and symbols. Do I have now in my hands just another meaningless content that stands in need of new rules in order to acquire full blooded sense? This would lead us to infinite new signs. Platonism says that what stops the regress is precisely the entity with which we get in contact when we understand a word or a sentence. But, if this entity is only known in such a strange and private way, it seems that Platonism does no more than postulate a "regress stopper" to fill the gap between a word and its interpretation. Third, how are we supposed to have knowledge of the Platonic entity that confers meaning on the signs? It seems that we should investigate a whole new metaphysics about things that are hidden to us without even knowing if our language means the same thing to each of us by the words "metaphysics" "of" "things" "hidden" "to" "us" . So, Kripke seems to reasonably dismiss Platonism as a possible answer to the skeptic.

The last possibility analyzed by Kripke before he presents the skeptical solution to the paradox is the dispositionalist answer. Dispositionalism claims that the interpretation of a rule consists in my reactions to it. So, whenever I want to know my interpretation of a rule, like a signpost, I should ask myself what would my reactions to it be. For example, if I see the sign  by the road, my disposition to stop settles the meaning of the sign, for it says no pedestrians are allowed. The advantage of dispositionalism is that it preserves the immediacy and directness that characterize our behavior in most cases when we are following a rule. However, there are some negative points about this explanation that will show why it is not the best answer.

Let's suppose that the interpretation of a rule is given by my disposition. But rules, as we know, hold for infinitely many cases. They have possible infinite applicability, nonetheless human mind is not able to grasp all applications at once, because of the finiteness of human life. Imagine that we were to add two incredibly giant numbers; numbers that it would take 140 years to calculate. In the realm of my dispositions there certainly isn't a disposition to add such gigantic numbers as I add "12 + 8", which is almost an automatic operation. So, if disposition

here means “the ability acquired after exercise and training”, I certainly don’t have such disposition to sum gigantic numbers, even though nobody will deny that there is a correct answer to the operation.

Another problem with the dispositional account is that not always my disposition will reflect my knowledge of how to correctly follow a rule. Imagine that someone says “Can you please add $45+55$?” to what I answer “Sorry, I’m late.” If my disposition is not to answer the question in its terms, then would my disposition still be a valid answer to that question? Or even worse: Imagine my disposition is to answer it wrongly. Instead of saying “ $45 + 55 = 100$ ” I say that the result is “90”. A dispositionalist would have to make it fit in his theory all the wrong answers I’m disposed to give. Also, and this is surprisingly mind-boggling, this account implies circularity, for any account of my dispositions will be given in terms of my dispositions to explain dispositionalism itself.

Of course, a good dispositionalist would try to save his theory saying that we should tell the good dispositions from the bad ones and shape our answer taking in consideration the good results. He will say that we can separate the ones that work and the ones that do not. The ones that work should be quantitatively and qualitatively superior to the bad ones, but this does not free us from the possibility of being massively mistaken. Just as for thousands of years our disposition to answer to the question “Where is the center of the universe?” was “the earth” and, for all I know, that was a massive mistake.

So, if my dispositions to answer a question in a certain way do not conduce me any closer to show any fact that could demonstrate that I mean “plus” instead of “quus”, we are allowed to dismiss dispositionalism and try to understand Kripke’s solution to the paradox. We have seen his three responses to the rule following paradox.

5.2. The skeptical solution

Kripke believes that Wittgenstein’s PI suggests a skeptical solution to the paradox. No fact has been alluded so as to derive from it our ability to interpret and understand meanings, but we still have to make sense of the fact that in our ordinary practices we constantly understand each other. Kripke will say that Wittgenstein “provides both conditions under which we are *justified in attributing concepts* to others and an account of the utility of this game in our lives.” (p.95) So, according to Kripke, what Wittgenstein is telling us is that, when I arrive at a room with people in it, I’m *justified in saying* “Hi!”, but also I could say “Good morning”. There’s no necessary next step or true next step to be taken when I arrive in a room, for all norms do for us is justifying us in attributing concepts to others. Hence, he is justified in saying “Hi”, because he just arrived, but he is also justified in saying “good morning”, because he just arrived. But the

problem is to understand how we constitute this notion of justifiability. Kripke will say that the key concepts to understand this practice are: agreement, form of life and criteria.

Agreement is reached when someone "exhibits sufficient conformity, under test circumstances, to the behavior of the community." (p.96). Evidently agreement can be gradual and can be (gradually) taught. It does not need to be total, but it must be enough so that even when there's disagreement the practice of giving reasons could make us understand the roots of the disagreement in question.

Also the agreement among parts brings about a further result. As far as our responses agree we constitute a *form of life*. We recognize each other as sharing the same reactions, either the spoken or the bodily ones, and for that reason we recognize others as sharing some very general features that we attribute to ourselves, like rationality and sensibility. In our case, people who follow some more or less consistent group of norms, behavior and emotions we call humans and when a person is educated this person is accepted into a community called humanity. As Wittgenstein says "what has to be accepted, the given, is - one might say - *forms of life*." (PI 305)

It is important to emphasize that sharing forms of life does not mean that we grasp the concepts in the same way, for meaning has not to do with psychological images or phenomena that happen to us when we think or talk, but, as Frege approximately would have said, meaning is a logical property of thought. So, if we belong to the same form of life, that means that we *can* share the logical features of thoughts, but it does not entail that we will have the same psychological experience upon hearing and uttering thoughts. What I am allowed to conclude is just that others legitimately use concepts in the same way I do. Why am I able to recognize this? The Wittgensteinian answer is: because the use of a concept manifests the criteria for its employment. Kripke will comment that Wittgenstein's idea that "Inner processes stand in need of outward criteria" helps understanding how we learn to apply psychological concepts, like the concept of pain. So, when someone thinks that my wince and my trembling are the evidence of pain itself, that means that the persons mastered bodily behavior as representative of pain - so this person sees outwardly what might be going on inwardly, but these criteria do not tell me if what I see is actually felt pain. It only indicates me that *this* behavior is *about* pain. The aboutness of a determinate behavior is grasped in the way it is manifested. Pain usually manifests itself by means of cries, contortion and wincing, so Wittgenstein called these manifestations as constituting the grammar of the phenomenon. Hence, the correct use of a concept manifests the grammar of its employment, and what we master when we learn to use concepts is the very general criteria of their application.

In this section I haven't been saying anything about the possibility of private languages, but I hope it's enough to point out the connection between the rule-following problem and the

private language problem. The possibility of private language is exactly what the skeptic about other minds needs to prove his point that we don't know the other's existence. If meaning is experienced in privacy, then we can't know what goes on in the other's mind. As I said earlier, a private language must be a language that make sense only to the person who's able to speak it and can't make sense to others. If we take the teaching of the rule-following problem as a guide to understand the private language possibility, we will see that the later argument also is supposed to demonstrate is that there are no facts that I can possibly present to myself to show that I am right in meaning x instead of y. As we have seen, Kripke's conclusion shows that meaning is somehow grounded in my sharing criteria, in my agreement with others and the form of life I share. Hence, for case of psychological phenomena like pain and feelings, the same holds: we can't mean them in privacy because we learn them by learning to read expressive behaviour, by imitating other people and learning whatever else is public and shared concept. Hence, I cannot in absence of any sense of otherness determine if I correctly felt the same sensation S twice.⁴

6. Wright and the Fact of Meaning

Wright rejects Kripke's understanding of Wittgenstein's rule-following argument. He will instead propose a so-called "flat-footed" answer to the skeptic, trying to meet his demand for a fact. Wright will say that it is open to view that *I remember that I meant* the word "plus" and not "quus". In his opinion this is a fact about me, based on my first personal authority, that the skeptic cannot dismiss.

Thus, Wright believes we can have knowledge of the correct application of rules, contrarily to Kripke's opinion. Knowing if we followed correctly the rules is a matter of tracking correctly our mental acts. As Wright himself says, the rule following problem "have nothing at all to do with the reality of rules. Rather, they are epistemological: Wittgenstein is preoccupied with the sense, if any, in which a rule is genuinely an object of intellection, something whose requirements we *track* [...] by grace of some intuitive or interpretative power." (239)

If he is right, our question is: How do I know if I have correctly followed a rule? Wright presents us a biconditional that summarizes his thoughts. "X intends P if and only if X is disposed to avow the intention that P, and would be sincere in so doing, and fully grasps the content of his intention, and is prey to no material of self-deception, and so on..." (WRIGHT, 2001b). This biconditional relates a set of epistemological conditions that must be met in order

⁴ Kripke, nonetheless alerts us to the the difference between following rules privately and following alone in physical isolation from others. Robinson Crusoe does follow rules when he is alone in an island, because he might be referring to what he has learnt previously in his community. It's not a problem the fact that he is alone. Even if he was born there alone and miraculously grew up among the trees, the mere acknowledgement of the world as a sufficiently regular sort of otherness would be enough, for it would play the role of a public support for his mind to follow rules.

to avow an intention and in the very act of avowing my intentions I obtain knowledge of my own intentions. In other words, our knowledge of the steps in following rules is given to us if we are able to retrospectively recognize our own intentions. But how do we recognize them?

Wright believes the biconditional can be read in two ways. The first way is a *detectivist* way, where the intention is first experienced and the subject is then able to *apprehend* it. A second way is the *constitutivist* way, where the avowal constitutes what the intention formerly was. Detectivism and Constitutivism are terms of criticism used by David Finkelstein in *Expression and the Inner* and I'll make use of them given their organizational and analytical powers. Finkelstein criticizes two forms of detectivism that he dubs as old detectivism and new detectivism. The old one relies in the belief that the mind is a medium to which each person has immediate privileged access. The classic old detectivist is Bertrand Russell. As he says in *The problems of Philosophy* (1912), "when I desire food, I may be aware of my desire for food; this 'my desiring food' is an object with which I am acquainted". So, when I think "I'm feeling hungry" I don't need to ask anyone else about what I feel, for I have immediate awareness of my own state. So, old detectivism takes one's relation to one's own mind as a reliable and privileged field for knowledge.

Nonetheless, old detectivism leads to problems. Even though it seems clear that no one is in a better position than I to say what I feel or need, understanding the mind as such a super reliable environment is certainly problematic. It pictures the mind as if no thought or feeling could be possibly missed or misinterpreted, but this picture is somewhat suspicious, since we visibly have many doubts during life about what feelings we have for people, or how we should understand situations and so. It is also problematic to think that I am always able to detect what I think or feel, since there are many experiences that I need to learn to detect. Finally, when we try to think of mind as such a reliable medium we don't know exactly if we should take it in a materialist or immaterialist conceptions. If we are to believe mind is such a stable and reliable environment, it is hardly a material one, since matter is unstable and prone to change. Also, it is remarkable that as we get old, we lose access to memories and we reason slowly. Hence materiality and reliability can be incompatible. But if materialism is a bad alternative in this case, the other solution will be to trust some form of immaterialism. Nonetheless, if the mind is immaterial, we fall in all the problems of dualism.

In opposition to old detectivism, new detectivism admits first person fallibility in knowing about my own states. In this view, my knowledge about my own mental states is inferential. The procedure to know my own experiences consists in observing myself with a sort of inner eye, to this I should add my knowledge about other people's reactions and my own body that my mental states is X, from this I should draw an inference. My fallibility resides in the fact that I know about my inner states because I draw inferences about myself and, since the nature of

inferential thought allows for error, I can be mistaken about the meaning of my own thoughts and feelings. New detectivism brings up a more natural conception of what it is to be in touch with one's own experiences, but a strange conclusion of this is that the way I know my mental states is on a par with my knowledge of external objects and that seems not to be the way I actually know many things about myself, e.g. when I feel pain I don't draw any inference.

Russell's later view is new detectivist. He says that "the discovery of my own motives can only be made by the same process by which we discover other people's, namely, the process of observing our actions and inferring the desire which could prompt them". (RUSSELL, 1921). This view unfairly takes the inner world and the outer world as equally accessible. Old detectivism had a point in taking our kind of awareness of mental states as immediate, but it was wrong it transforming it in a sort of undoubtful and privileged form of awareness. New detectivism mistakenly goes to the other extreme. It seems that it has forgotten that our minds give us some sort of authority in accessing our own states that need not be inferred from our behavior - as if we were an audience that observe our own selves.⁵

6.1. Wright's Constitutivism

So, back to Wright's biconditional, the only option left is the *constitutivist* reading. Wright believes constitutivism is a sort of decision, a sort of agential process of the mind in "making meaning". Wright also believes this is Wittgenstein's view, for PI 186 says that "Intuition can mislead you. So, a new decision is needed at every stage". Intuitions are knowable by detection. They are primitive, unarticulated in their apprehension, and a form of knowledge too basic and immediate to admit of any further account. If the process by which we identify new steps in following rules is by intuiting the new steps, then we are led to adopt the detectivist account, but then, we have the same problems faced by the detectivist account. Hence, new steps in rule following are not intuited, they might be constituted. Constitutivism suggests a subject actively doing his own awareness of the on goings of his mind - deciding what the next step is in following a rule. This is why to follow a rule is a certain kind of decision in this view. But, of course, although it is a kind of decision, it is not a kind of decision with no objectivity. Wright thinks that our best decisions are made under a regular set of conditions that enable us to know when we are correctly following rules - that is the importance of his epistemological point.

In a nutshell the core argument of Constitutivism is that I make it the case now that what I said before is what I am now saying. "The fact about my past usage of "plus" that fixes it that I am now acting in accord with what I then meant by "plus" is just that I meant plus by "plus" says

⁵ In *Expression and the Inner* by David Finkelstein those arguments can be seen in much more detail and the criticism is also expressed in a much more refined way. Here I just want to state their incapacity to solve the problem of meaning regarding such we are trying to understand if the concept of mind has public acceptance.

Finkelstein (p.39) That means that I stipulate what the rule requires by retroactively conferring the right meaning to it, then "every rule and every intentional state get its content by a kind of stipulation."

Now, even though Wright's solution frees us from Kripke's "skeptical solution", something is too strong in his *constitutivist* view. The separation between the first mental act (such as perceiving, thinking, etc) and the second one (the constitutive act) makes Wright's solution the constant practice of bridging an act deprived of meaning and another act that will confer meaning to it. It is as if communication and language were the action of giving dead signs a sort of vivacity -- a dead sign acquires its meaning by means of an interpretation that constitutes its full blooded sense. This, in Wright's view, is consistent with Wittgenstein's interlocutor at PI§ 431: "There is a gulf between an order and its execution. It has to be filled by the act of understanding." So, there's a separation between my understanding of the rule and the rule itself. But then, what is the relation between my interpretation and the sign? How do I connect both? If I ever see a sign, I will have to interpret it and my interpretation will be expressed by means of the signs I use to express my thoughts. So, I'll again have to interpret the signs I used to interpret the first sign, which will raise a third level of signs to be interpreted. Of course, this will go to infinity and we never finally interpret the first signs that were staring in front of us.

Hence, Wright's idea of stipulation is an attempt to answer the rule following problem with a controlled and methodological way of giving meaning to dead signs. He wants the dead signs to be filled with meaningful content by an "act of understanding" where our best opinions are taken into consideration at the moment of stipulating the meaning of a word. Now, looking back at Kripke's paradox we see how the fact that "no course of action can be determined by rule, because every course of action can be made out to accord with the rule" (PI§201) is connected with the constitutivist account of rule-following. Given that at each step I constitute the proper meaning of the rule, I would have solved the paradox by appealing to this fact. However, as it was said above, this solution entails a regress where we can never bridge the gap between the symbol and its meaning.⁶

Now we finally can establish an important connection to our main subject. If constitutivism was right, then the bodily movements would have to be seen as any other sign. Running, clapping hands, bowing, etc are meaningful bodily attitudes. So, the same account

⁶ In pages 45 and 46 of *Expression and the Inner*, Finkelstein will also explore another sort of objection. Although Wright's constitutivism focus mainly on attitudes, Finkelstein extends the constitutivist claim to mental activity. The responsibility objection to constitutivist account blames constitutivism for excessively attributes first person responsibility in the determination of meaning. It is as if whenever I speak about my headaches, I'm actually making the case the I have one. So, if my headaches were constituted by my self-ascription and avowal, then it would make sense to blame me for them. We know that not all headaches come from hangovers (which would be the case to blame me as responsible for the headache), but we know that some are involuntary. Constitutivism would attribute even the involuntary ones to the agent. So, constitutivism puts too much weight on the responsibility for one's own mental states.

applies: whenever I interact with another person, this person would be seen as a body or a behavior to be interpreted; but as Finkelstein puts it

If we have been moved to think that there is a gulf between every rule and what it requires, we will also think that there is a gulf between any human behavior and its psychological significance. It will seem to us that what we perceive when we look at another person are bodily movements and their effects (including sounds made by speaking), and that some sort of interpretation of these movements and effects is needed if we are to learn about the other person's intentions, wishes, hopes, etc. (Finkelstein, 2003. p.43)

Constitutivism is consistent with the common view on the problem of other minds on which the other's body is taken as the raw material to be interpreted. So, if we were led to this point in our connections between the rule following case and the problem of other minds, it seems that we can expect an answer that can solve both the rule-following problem and the problem of other minds at one stroke.

7. McDowell and Disjunctivism

Now we shall summarize what the problems with old and new detectivism as well as with constitutivism. Finkelstein will classify those three attempts of solution according to three features that might correctly explain my access to my own thoughts: intimacy, responsibility and naturalness.

First, intimacy is the property that I have in virtue of my privileged access to my own thoughts. My thoughts are mine and my first person perspective on them cannot be overlooked when we are trying to explain the activity of meaning.

Secondly, responsibility is the property that I have as the author of my own thoughts and as the owner of a mental life. Responsibility is essential to understand our mental activity, for whatever happens to me in the realm of my thoughts, sensations and feelings belongs to me and, to a certain extent, can be traced back both to me as the author and owner.

Naturalness is the property of thoughts if they are to be mental events connected to the reality of my body and the objects. Any natural account of my mental life and meaningful activity cannot see my body in the strangeness of an eye disconnected from the material world. As natural creatures our mental life cannot be depicted as a realm of thoughts freely floating in an ethereal existence. Pain, for exemple, is something that evinces the connections of mental activity and Our mental life must be connected to the kind of natural creature that we are.

Old detectivism is strong on intimacy and responsibility, but weak on naturalness, given that it relies on a super mechanism of detection of my own thoughts. New detectivism is strong on naturalness and responsibility, but weak on intimacy for my access to my own thoughts is the same that I have to the thoughts of others. And constitutivism is strong on intimacy and naturalness, but weak on accounting for responsibility, given that my responsibility over mental events is overestimated. Finkelstein tells us that we would like an approach that is strong on all those three features. His suggestion will be

Hold on to what's appealing about constitutivism - the idea that we can respect both intimacy and naturalness by introducing a constitutive moment into our story about the relations a subject bears to his own inner states - without committing ourselves to the problematic claim that in order for someone to speak with authority about his own state of mind, he must be active in its constitution. [...] Is there a way to say both (1) that my awareness of, e.g., my current headache in some sense constitutes it *and* (2) that the constitution of my headache is not due to my activity, so not something for which I am responsible? (Finkelstein, 2003, p.53)

McDowell's books from 1996 and 1998, respectively *Minds and World* and *Meaning, Value and Reality*, will argue against Wright's constitutivism. His claim is that Wright's argument is somehow embedded in a long tradition that stretches back to a modern anxiety. The way modernity construed the concept of mind opened a gap between mind and world in such a way that all the efforts from Descartes to contemporary philosophy have been to conquer back the lost unity. McDowell does not want to bridge back the gap from mind to world or from world to mind, he wants to show us that no gap was ever opened.

McDowell begins with a very simple idea: sensory experience constrains thought and judgment. So, if I *see* that it's sunny outside I can also *judge* that it's sunny outside. I could imagine that it is not sunny outside, but if do so aware that it is just an operation of my mind, I won't believe my imagination. Perception constrains me to believe and think that it is. In Kantian terms, we are sensible beings that have the receptive capacity to be affected by external objects. Receptivity constrains our understanding that spontaneously articulates the proper concepts to answer to our sensibility.

This means that sensory experience plays a role as epistemic ground for judging and thinking about the world. This view is in disagreement with Wright's, for he distinguishes epistemically between objects in the world and the way we understand them -- the constitutivist view confers meaning to experience only in a second moment when the subject is capable of

avowing his/her first experience. The idea that experience is constituted in two logically separate moments entails a twofold understanding of experience; one of meaningless brute experience, another where meaning is constituted with the help of the understanding. He believes we can have non-conceptual sensory experience of objects (a sort of independent experience of given raw sense data), as if we experienced first brute sense data, and then constituted the object and the meaning of it in experience. Hence, Wright claims that there is a given in experience that affects me independently of my intellectual powers.

In this debate, the term widely used to talk about sense data was introduced by Sellars (1997)⁷. He calls sense data “the *given*”, implying that the nature of our receptive powers is passively affected by sense data from the outer world. The common recourse that philosophers have made to explain perceptual knowledge allows Sellars to identify a common “Myth of the Given” as a sort of foundationism for epistemology. He denounces the epistemology that has fallen into the Myth of the Given as a theory of knowledge that tries to build up from the bare sense data the explanation of all perceptual knowledge. The problem of falling in the Myth is that we face the same problems Wright faces with his constitutive conception -- also known as additive conception. We stand in need of something to be added to these brute data. McDowell on his side, will acknowledge the need of a minimal empiricism, for we need to make sense of mental states episodes that are directed towards the world.

McDowell avoids the sense data conception of the given, but he also wants to avoid falling into Coherentism - the belief that all my knowledge comes only from logical relations among beliefs and justifiability relations. If all knowledge comes from logical relations reflecting coherence among beliefs, we could have empirical knowledge where our sensory experience does not ground any sort of knowledge.⁸ *Mind and World*, as Finkelstein dubs, is a middle path between those two radical conceptions. McDowell does not want our knowledge of the world to be the addition of meaning to brute sensations, but he also does not want our knowledge to be a “frictionless spinning in a void” that scoops out sensibility from our concept of knowledge. In the middle path account “conceptual capacities, capacities that belong to spontaneity, are already at work in experiences themselves, not just in judgments based on them” (1996, p.24)

The defense of the claim that our sensory experience constrains our conceptual capacities makes appeal to the Sellarsian concept of a “logical space of reasons”⁹. If my perceptual experience did not passively actualize my conceptual capacities, my conceptual

⁷ Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." *Minnesota studies in the philosophy of science* 1.19 (1956): 253-329.

⁸ Coherentism here is a label applied to Donald Davidson's conception of knowledge (1986)

⁹ “The essential point is that in characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says” (EPM: §36)

experience could not ever be situated in a logical space of reasons. Therefore, I would never claim to have knowledge of something that I have seen, heard, etc. That is, I can situate my experiences in a chain of reasonings, if it's necessary. For example, I can walk down a street without making any judgments, but when I arrive at the place where I'm going to meet my friend Josh, he asks me if I noticed that there was a new Walmart in the corner of 63rd and Ellis ave. Given that I have seen that, I can place it in the logical space of reasons as a perceptual experience associated with my recent memory. Now, imagine my friend insists that if I walked down the street I should have seen the new supermarket. But suppose I didn't see it, even though I walked the street and I know I should have seen (the inauguration was that day). However, if I can't place the experience of seeing the supermarket in the space of reasons, this means that even though my perceptions are causally explained to certain extent (e.g., light penetrating my eye belongs to the logical space of nature), my *experience* of nature is not causally determined, for it depends on the passive actualization of concepts only possible given the spontaneity of mind.¹⁰

Now, what about the concepts of our mental life? How should we think them? McDowell asks us to think of the inner realm in a similar way we think of the outer experience.

The realm of thought and judgment includes judgements about the thinker's own perceptions, thoughts, sensations and the like. The conceptual capacities that are operative in such judgments must belong to spontaneity just as much as any other conceptual capacities do, and that can generate the spectre of a frictionless spinning in a void for this region of thought too. Then, in the way that should by now be familiar, ensuring friction, which is required for genuine content, can seem to oblige us to take exercises of concepts in this region to be rationally grounded in something extra-conceptual, bare presences that are the ultimate grounds of judgments. (McDowell, 1996, p.18)

McDowell holds that in the same way the outer world constrains our conceptual capacities, the inner world also does, but in this case the friction is the other. The other's body, the other's reactions and words give outer criteria by means of which I learn to conceptualize my inner world. As Wittgenstein says in PI 580§, an 'inner process' stand in need of outward criteria. But what is this need about? I think inner processes need outward criteria in order to

¹⁰ A common objection to McDowell's conceptualism brings the possibility of experiences that cannot be conceptualized. We can think of shades of colors that are unknown to us and have never been conceptualized. McDowell's answers to this objection by reminding us that we don't need the actual specific concept to have the experience of the color. What is necessary to spontaneity of thought is just the minimal conditions of ostensive thought, e.g., a "that thing" or "that shade" held in thought.

have nothing less than meaning. If there weren't any public observable manifestation of the inner, in others and in me, we would never agree in our reactions. We would never know what being in conformity to someone actually means. We would not know what it means to express yourselves otherwise than linguistically.

Criteria give us the indication that some behavior is about pain, or about terror, or about being in love -- the experience of fiction is possible because of this kind of notion. Criteria does so because they are connected to the grammar of phenomena. If wincing, contorting and screaming are expressions of the phenomenon of pain, then certain sounds, outer behavior and words constitute the grammar we use to talk about pain. Accordingly, criteria do not give us certainty about what goes on in the other's mind, all we can know is that this person's behavior is about, e.g., pain. So, sharing criteria is fundamental to constitute our reactions, for it is by means of criteria that we express in meaningful ways our inner lives.

Other persons as well as my own body release me criteria that allow the identification of mental phenomena. Inner awareness, differently from objects of perception, doesn't have an existence independent of the mental experience that constitute it.

It may seem that we are back with constitutivism about our own mental events, but McDowell firmly rejects Wright's additive conception. He's not saying that the existence of pain, feelings and thoughts depends on my avowing and speculation of any former intention. Criteria allows us to recognize mental phenomena, which is different from simply creating them. Obviously recognition of one's own mental environment moves certain things from darkness to light, which may provoke the feeling of bringing something to existence. However, the movement is actually one of domesticating and civilizing raw nature and shape it into shareable human-like expression. Pains, thoughts and feelings exist in the way consciousness manifests them: bringing and presenting objects that exist in the form of awareness of themselves. McDowell isn't a detectivist, for even though the mind also is understood as an intimate place, my knowledge of my thoughts is achieved, not by means of any supernatural mechanism, but because there's no distinction between being aware of a thought and the thought itself. As Finkelstein puts it: thoughts, pains and feelings are constituted by the awareness of them.

You might be wondering: the rule following discussion lead us pretty far from our initial intention of knowing whether the concept of the mind is public. I did not lose track of that. Our discussion about rule following has the intention of showing that if meaning is to be rule governed, then it must be possible to follow rules. Now, I'll explain why it is possible to follow rules, and given that we bet it is possible to follow rules, then to make sense of my own mental life as well as of the mental life of others is a special case of following rules.

Now I want to account for McDowell's understanding of the rule following problem. We can envisage it from his conception of experience and knowledge. His view on knowledge and

perceptual experience tie mind and world in double aspect unity. The world we experience is the experience of a mind within the world that conceptualizes this independent world. The mind is not outside the world or in the limits of the world observing reality, as the *Tractatus* has depicted. His view is known as *disjunctivist* for it takes our experiences to be a disjunction of either the experience of x or of y, in the following sense: my perceptual experience of a brown chair is *not* the *conjunction* of factors like brown shades of color (spatio-temporal distributed sense data) *and* the concept of chair (and of unity, substance, objects, etc). This conception is precisely the one involved in Wright's understanding of rule-following, since he takes signs to be meaningless data standing in need of acts of mind. According to McDowell, if my perceptual experience is of an actual chair, then the actual chair is given as a whole (perceptions, concepts and all included) or it is not given in actuality at all.

A classic argument that is based in conjunctivism is the skeptical argument from illusion. This argument is designed to state the existence of sense-data. A visual illusion, for example, deceives a person who believes to be perceiving an object. Someone who sees a half stick submerged in water may have the illusion that the stick is bent. Ayer¹¹ (1940) generalizes this possibility trying to show that we are never directly aware of any objects in perception. All we have access to are sense-data. However, illusion can be radical, like the cartesian dream scenario, or can be local and specific, as Dretske's zebra-mule example. Both scenarios involve the experience of something illusory taken to be real; the first states we *never* know if what we experience is real, the second example states that we *cannot* know if the *whole* experience is real.

As Cavell has demonstrated in the Claim of Reason, the argument from illusion triggers a search for rational principles that can guarantee the existence of the world. The argument from illusion starts challenging our capacity to know if the ordinary objects to which we are presented are really what we perceptually experience. Given our limited perspective on the world, we are never in a position to affirm, with certainty, to have knowledge beyond our immediate experience. To give you an example, I never know if the apple I see on the table isn't just another slice of a fake reality that self-destructs after my experience of it, for whatever proofs I might empirically construct would be under the same suspicion of fakeness. The nature of our knowledge of the existence of an external world is inductive and limited, which is insufficient for a skeptical demand. The skeptic will force you to see that, if you cannot be sure about the existence of an apple, then the knowledge about the whole reality is equally threatened. His criteria of knowledge as certainty starts with the simple act of perception and reaches a degree of maximal generality. The doubt does not extend to only this world and this reality, but to whatever externality to my own mind I can imagine. The skeptic's criteria, from the moment it

¹¹ Ayer, A.J. *The Foundations of Empirical Knowledge*. (1940)

enters the argument, forces us to accept that the analysis of experience has the structure Wright has conjectured: We cannot say whether we have access to objects in themselves, what we can undeniably say is that we have the experience of shades of color that help us compose something called experience of the world.

Wright's constitutivist model of experience is modern in spirit. It analyses experience to its ultimate basic components and mistakenly equates different forms of experience due to their basic components. So, given the analysis, illusion and reality share basic common features as colorfulness and objecthood, in the words of a skeptic. The difference between one and the other rests in the way we mentally construct reality or dream out of the common features, but no longer in the objective value of experience. Therefore, we may not be able to distinguish illusion from reality.

The mistake in this argument, in McDowell's jargon, lies in describing experiences taking the myth of the given as starting point. The argument claims that both perceptual experience and perceptual illusion share a highest common factor (sense data) and, in order to determine whether we are really seeing an object or merely having an illusory experience of it, something must be added to the given. McDowell's disjunctivism, as earlier described, entails the rejection of such a conception, because the unity of perceptual experience does not require anything to be added to it in order to become knowledge or awareness. Perception is metaphysically distinct from illusion, so the analysis of the knowledge of each one must start from the distinction instead of presenting it as a result.

The same happens with rules. McDowell believes that "Wittgenstein's point is that this dilemma [raised by Kripke] seems compulsory only on the assumption that understanding is always interpretation"; actually Wittgenstein's intention "is not to shift us from one horn of the dilemma to the other, but to persuade us to reject the dilemma by discarding the assumption on which it depends" (McDowell, 1998, p.238).

Why must we read Wittgenstein this way? McDowell alerts us to read §201 until the end and to carefully consider §202.

That there is a misunderstanding here is shown by the mere fact that in this chain of reasoning we place one interpretation behind another, as if each one contented us at least for a moment, until we thought of yet another lying behind it. For what we thereby show is that there is a way of grasping a rule which is *not* an interpretation, but which, from case to case of application, is exhibited in what we call "following the rule" and "going against it"

That's why there is an inclination to say: every action according to a rule is an interpretation. But one should speak of interpretation only when one expression of a rule is substituted for another.

That's why 'following a rule' is a practice. (PI 201-2)

The end of §201 entails that the rule following problem can be read in a disjunctivist way. Rules, as Finkelstein explains, should be situated within the stream of our lives - in our practices as 202 suggests. Obeying a rule cannot be the same as interpreting it. Interpretation is the kind of thing we need when (e.g.) we *don't understand* a police officer telling us to stop in a foreign country. As McDowell says

I think the thesis that obeying a rule is a practice is meant to constitute the answer to this question. That is, what mediates the inference ("hence also") is this thought: we have to realize that obeying a rule is a practice if we are to find it intelligible that there is a way of grasping a rule which is not an interpretation. (McDowell, 1998, p.238)

Gulfs between rules and my capacity to follow them might appear when I am presented to new rules, new situations and contexts, and these are the cases in which I might need an interpretation; those are the cases in which I might need to learn how to follow the rule. We learn to obey rules as part of our education, but to get to it we rather enter in agreement with human institutions, customs and practices that precede us. This is not to say that what we should add to the rules in order for them to make sense are institutions, our history or our community, but it is to say that they only make complete sense *seen from within* the perspective of a shared form of life.

Back to the concept of the human mind, we can say that 'mind' is a public concept, because it is the concept we constituted in order to talk about expressive human behavior. Human institutions educate bodies and behavior. They teach us when to manifest affection, which are the feelings we should direct to whom and what relations we establish in each context. So, McDowell's answer to the rule-following also serves the purpose of understanding how we relate to other minds, since rules determine and teach us forms of bodily behavior. Another consequence of this educational target is the capacity to develop the awareness of our own selves, since the experience and the conceptualization of our inner lives leads to the awareness of oneself as a self-conscious being. Following PI§580, McDowell agrees that the practice of conceptualization of our own outer behavior and of other's bodily expression maps

the manifestations of our inner life, and this gives us criteria to grasp the concept of mind as the concept that represents our inner world. This might not give us yet knowledge of other minds, but it points to the possibility of sharing the meaning of what is to be the "owner of an inner life".

Chapter 2

From the Concept of Other Minds to Acknowledgment

The human body is the best picture of the human soul.

Wittgenstein, Philosophical Investigations

1. Summary

The discussion that started with Kripke and led to consider McDowell's view on the rule following problem will enable us to better understand the conceptual problem of other minds. The concept of other minds is only possible because the normativity underlying our conceptual capacities is grounded in our practices, institutions and forms of life. I don't mean this answer to commit us to another sort of foundationalist conception of rules, rather, what it teaches us is that the objection that I cannot apply the concept of a mind to myself in the same way I apply the concept of mind to others is groundless.

From the fact that the concept of mind is applicable not only to me, but also to others, it does not follow that the application gives me *knowledge* of other minds. This possibility can still fail, so we haven't got rid of the skeptical threat. The publicity of meaning entails that I *can* know, but not that I *actually* know of the existence of other minds when I apply the concept. In what follows we will focus on attempts to account for acquisition of the concept of other minds.

The first attempt to argue for the possibility of knowing other minds that is not analogical, but still inferential in nature is Theory Theory account. Theory Theory denies the asymmetry between my mind and the other's, for arguably I don't know my mind in any special way. The idea is that I learn the concept of (my) mind in the same way I know yours, given that I know mine by means of the same concepts and no first person privileged access is achieved without the help of these shared concepts.

From the failure of Theory Theory we will see a new and promising theory. Interaction Theory recently enriched with findings in psychology, neuroscience and anthropology gave birth to Enactivism. This theory claims that our capacities are to be understood as a system of body-brain-environment that allows us to redescribe our cognitive dynamics.

The description of our cognitive dynamics will be followed by Strawson's arguments concerning other minds. He claims the concept of person to be a primitive unity. His theory

explores the notion of criteria and the conceptualisation of first personal and third personal mental life. With it he will deliver us a better understanding of the dialectics present in the acquisition of the concept of person.

Before we reach the end, I will briefly explore a different way of seeing things. Nagel's reading of Sartre's *Cogito of the Other* is certainly an interesting point of view on how we get to know ourselves as objects in the look of the other. Sartre thinks our awareness of other minds is first enacted as the awareness of being observed by someone, as in a passive position.

Finally, the end of this chapter will reconstruct Cavell's analysis of the skeptical problem of other minds. His analysis is centred in the grammar that structures the skeptic's and the anti-skeptic's point-of-view. Cavell shows that both are captive of two different images. The skeptic is prisoner of a first-personal standard of certainty. His opponent is caged in a merely third-personal standard objecthood. Both criteria for knowing other minds are insufficient. Cavell will propose that the problem of other minds is motivated by the lack of acknowledgement of the other. As if a sort of blindness to the humanity of the other was at the root of the problem.

2. Theory Theory

Paul Churchland in *Scientific Realism and the Plasticity of Mind* (1989) claims a version of Theory Theory. Theory theorists believe our common sensical notions from folk-psychology are a form of proto-theory with which we learn to interact with ourselves and others. Churchland argues that "our common sense conceptual framework for empirical reality is in all relevant aspects a theoretical framework." (p.89) Thus, according to his view, a certain general *theory of mind* plays a role in our ordinary interplay with other minds (and also in my experience of my own mind). Churchland claims that arguments from analogy are problematic because they are "essentially parasitic on one's knowledge of one's own mind" (p.90-1). So what we must follow instead is our common sensical theory of subjectivity, because it is an unbiased general theory. The evidence that we have for this theory is that we have a set of sentences that describe not only types of human behavior, but also types of inner states and the causal connections between both.

Consider again the set of generalizations whose justification is at issue – the set of sentences descriptive of the general relations holding between (a) types of causal circumstances and types of psychological states (hereafter: P-states), (b) the various types of P-states themselves, and (c) types of P-states and types of overt behavior. We need only think of this set of general statements as a *theory* of the inner dynamics of human beings, ... as a theory whose credibility is a direct

function of how well it allows us to explain and predict the continuing behavior of individual human beings. If its prowess in these respects proves considerable, then one has paradigmatically good reason for accepting that dynamical theory as true” (Churchland, 1989 .p.91).

Churchland believes that we build up a theory to understand human behavior in general. This theory will give me a dynamic description of a human mind in terms of types, and given its *generality* it will be possible to understand both me and others without needing to start from my own case. Accordingly, by means of these descriptions we reach a set of detailed hypotheses concerning the types of determinants of human behavior, and from this set of hypotheses we can predict our own behavior and inferentially know what a human mind looks like.

The other feature of his theory, in addition to generality, is independency.

[O]ne would have these reasons, note, *independently* of any appeal to the facts of one’s own case. Conceivably, the facts of one’s own case might even be very *different* from what the theory asserts (one is a Martian, say, with a radically alien psychology), but one could still have excellent grounds for embracing it as true of humans generally. In principle, then, one’s epistemic access to the minds of others need nothing to one’s access (if any) to one’s own (idem. p.91).

Churchland adds that it's not only because of its generality that Theory Theory is a good way to obtain knowledge of other minds, but also because the status of this theory is independent of my own case. If the theory enables one to draw the correct predictions about human mind, then it has meaning independently of the kind of psychological environment any particular person was constituted. That's why the the dynamic theory of other minds can justify my knowledge of my own mind as well as the knowledge of the minds of others. Since this theory is *independent* of any knowledge that I have about my own case, but at the same time applies to me, I have a sort of inferentially tacit knowledge of others and of myself.

To wrap this up, Churchland's conception of the problem of other minds is that we can positively attribute knowledge about other minds because we are beings who have an independent and general theoretical standpoint about ourselves as a species. He says that as we develop a theory about human behavior and mind, and as it shows itself successful, we end up building a general Person-Theory of Humans.

The problem with Theory Theory accounts is that they seem to treat my own mental life as the life of a stranger. It puts all processes of knowing my mind and the mind of other at a personal level. Other than that, much evidence has been presented recently showing that "infants as young as 13 months have been shown capable of understanding actions, intentions, and, purportedly, beliefs (Onishi and Baillargeon 2005)¹². Hence, at subpersonal level and making no use of folk-psychology or any proto-theoretical notions. What better explains this form of scientific discovery is the Interaction Theory.

3. Interaction Theory

Strong criticism to Theory Theory comes from the so-called Interaction Theory. A good example of this progress has been presented by Gallagher and Somogy¹³ in their article called *Social Constraints on the Direct Perception of Emotions and Intentions (2014)*. There they "defend the idea that social understanding is primarily based on embodied social interaction" so that the knowledge of other minds is obtained by a sort of direct perception of the other's expressive (thick) behavior.

Behavior, in an embodied notion of other minds is not merely *surface behavior*. In an embodied account of our mental life, behavior is the expression intentional and emotional states developed by embodied beings with the capacity for responsiveness to their environment as much as the capacity to act on it. This capacity is called enactive, it works at a sub-personal level in the early life of humans and it is at the basis of personal capacities.

Gallagher and Somogy claim direct perception is a possibility in this case because for embodied-mind theorists intentions and emotions are not phenomena "in the mind". They are phenomena we early in life learn how to deal with at a subpersonal level.

Bodily or motor intentionality involves the aspects of meaningful motor behavior and expression that constitute what we call intentions. This idea attempts to capture the fact that the experiencing agent is intentionally engaged with the world through actions and projects that are not reducible to simple internal mental states, but involve an intentionality that is motoric and bodily. Actions have intentionality because they are directed at some goal or project, and this is something that we can see in the actions of others. (Gallagher, S & Somogy, V. 2014. p. 188)

¹² Onishi KH, Baillargeon R. Do 15-Month-Old Infants Understand False Beliefs? *Science (New York, N.y*

¹³ Gallagher, Shaun & Varga, Somogy (2014). *Social Constraints on the Direct Perception of Emotions and Intentions*.

So the idea is that an embodied-mind is engaged with the world and with others in a systematic unity of possible interactions. The embodied-mind is enacted by the capacity to interact with other humans so as to perceive them as expressing intentions. Hence, the nature of this interaction is not inferential in most cases. Obviously, depending on the context, the other's actions and intentions can be taken as meaningless movements waiting for meaning attribution, but this is the case where I haven't been trained or my experiences are not constituted so as to interpret the behavior in question.

The idea is that intentions at present (where I am able to specify what are my intentions in the action situation) or motor intention (intentions that intrinsic to make up my action) are not hidden. We can perceive these two forms of intention "without need for extra-perceptual cognitive inferences", since we constitute patterns of emotions by means of bodily-experiential-expressive aspects that are immediately perceived.

The perceiver is enactively engaged in perceiving the intentions of others, in such a way that her own motor intentionality contributes to perception. (idem. p.190)

Enactivism will allow us to think emotions as a pattern constituted by a larger system of brain-body-environment group of responses and reactions. This enables us to perceive them directly, with no further need of inferences. Theory Theory picture humans as inferential beings, who are in very subtle and unconscious ways thinking, but an enactivist description of the cognitive dynamic of other minds shows that there's no reason to suppose inferences where they are neither phenomenologically present, nor logically required.

This theory opens up a space for Anthropology and Cultural Criticism. When the enactivist account calls for environmental determination, when it calls for the well known Gibsonian concept of affordances (that determines opportunities for action), then we should also remember Social Affordances. These determine not only how we react to the environment at large, but also how we are to react to other's gestures, emotions and actions. Now we will see how the very general fabric of culture develops. What I'm calling the culture's fabric is what Strawson and others will call criteria -- the notion that expresses the agreement necessary for a concept to apprehend something in its raw form and find a place in culture for it.

4. Strawson and Criteria for Others

There is an asymmetry between the first and the third person psychological predicates ascription. However, this asymmetry does not need to be a two-step process, since even though

I ascribe my sensations (e.g., of pain) to myself necessarily in a first personal standpoint, it does not mean that I constituted them *uniquely* by means of my first personal standpoint. Strawson claims that I can only have the concepts of thoughts, feelings and sensations insofar as I can also ascribe them to others.

Strawson in *Individuals* holds that

One can ascribe states of consciousness to oneself only if one can ascribe them to others; one can ascribe them to others only if one can identify other subjects of experience; one cannot identify others if one can identify them *only* as subjects of experience, possessors of states of consciousness (Strawson, 2002, 100)

Strawson understands that it is a condition of possibility of ascribing states of consciousness to myself that I can ascribe them to others. This means that if I have criteria to discern my mind, I have criteria to discern other minds. For, if I can't understand what it is for someone else to have thoughts and feelings, then it makes no sense to speak of "my thoughts and feelings", for I would not be able to separate what is properly mine and what belongs to others. That is to say that the self-ascription of any inner experiences is only intelligible through contrast to what is "not-ascribable to me".

So, given that according to Strawson, we can only ascribe consciousness to other beings if we are able to identify other subjects of experience. Hence, I must be able to identify in reality beings that think, feel and have all sorts of conscious experiences. But it's reasonable to ask "how do I do that?". As we have seen before, any inferential identification of other subjects of experience takes bodily expression as criterion for ascription of mental content, but it fails when it leaps in the intimacy of the other's mind without warranted identification of the other as being a subject of experience. Strawsonian anti-reductivism claims the irreducibility of the concept of a person, for this very concept does not identify others *only* as possessors of states of consciousness. The concept of a person expresses a primitive unity.

What we have to acknowledge... is the *primitiveness* of the concept of a person... [T]he concept of a person is the concept of a type of entity such that *both* predicates ascribing states of consciousness and predicates ascribing corporeal characteristics, a physical situation, etc. are equally applicable to a single individual of that single type. And what I mean by saying that this concept is primitive... [is] that a necessary condition of states of consciousness being ascribed at all is that they

should be ascribed to the *very same things* as certain corporeal characteristics, a physical situation, etc.(idem, 101-102).

A bodily self-conscious being is a being that has the ability to express his own nature to himself knowing that the description of its nature makes use of both bodily criteria and physical situation correlatively to inner states. Both go necessarily together. Strawson wants to reject the Cartesian conception that we are a compound of two kinds of subjects: a subject of experiences (a consciousness) on the one hand, and a subject of corporeal attributes on the other. So, the way he kicks Cartesianism out of the scene is by claiming the connection of the criteria for conceptualizing the inner and outer.

We can say that *Individuals* claims that the inner and the outer are "spaces" that can only make sense if there's an inner experience and a body that I can possibly call "mine" in the same sense. For, if when I say "my body" and "my consciousness" I use two senses of "my", then the criteria for the unity of a person would be destroyed. Both uses must entail the unity of a person. Therefore, Strawson claims that the very idea of experience ownership comes from the possibility of calling my consciousness and my body mine or not-mine by means of the same criteria.

Thus, in order to ascribe experiences to myself, it must be possible to know that these are not inner experiences of another and that this is not the body of another. And it is not the experience of any other person because the conditions of possibility of calling something mine demands that I know that the same something should not be called another's.

It follows that criteria to ascribe mental properties to me as well as criteria to ascribe them to others must be public and open to view. These criteria are identifiable in the very same things that look like persons, that look like me. Now, we have the elements to characterize Strawson's opinion: why is the concept of a person primitive or irreducible? Because the simple possibility of ascribing this concept to something entails that another being must have both non-separately corporeal characteristics and inner states.

So, how is this supposed to help us answering the problem of other minds? How are criteria supposed to give me knowledge of the existence of others? All answers that make appeal to criteria bring up the problem of how we acquire criteria for something. In the case of learning criteria to discern my mental life, the case seems a special one for Strawson: he holds that the learning of criteria that apply to my own case is inseparably connected to learning criteria that apply to others, given the primitivity of both criteria. Strawson says that

“[J]ust as there is not (in general) one primary process of learning, or teaching oneself, an inner private meaning for predicates of this class,

then another process of learning to apply such predicates to others on the strength of a correlation, noted in one's own case, with certain forms of behavior, so—and equally—there is not (in general) one primary process of learning to apply such predicates to others on the strength of behavior criteria, and then another process of acquiring the secondary technique of exhibiting a new form of behavior, viz., first-person P-utterances. Both of these pictures are refusals to acknowledge the unique logical character of the predicates concerned" (Idem, p.107-108).

But if primitivity entails that criteria for self-ascriptions and other-ascriptions are learnt inseparably, without neither giving me any kind of priority over others, nor giving others any priority over me, then, in Strawson's view, "one must acknowledge that there is a kind of predicate which is unambiguously and adequately ascribable *both* on the basis of *observation* of the subject of the predicate *and* not on the basis (independently of observation of the subject)" (p.108). This is the case, to take one of his examples, of depression. If we imagine someone feeling/showing depression as having the structure of P(x), then when I fill the function with the first person "I", the sentence "I am depressed" this will tend to be understood in terms of "I feel depression", meanwhile when I put the third person, "he is depressed", the sentence is commonly understood as "he shows depressed behavior". This reflects the idea that "feelings can be felt but not observed, and behaviour can be observed but not felt" which Strawson takes to be deceitful, for we can have the illusion that we are not talking about the same thing. However, he claims that the structure of our language includes this perspective about things that are, in a perspective, observed but not felt; and in another perspective felt, but not observed. This is exactly what the publicity of criteria makes for us.¹⁴ They allow us to speak of persons as inseparable and primitive unities of mind and body.

This shift of perspective contained in language will be fundamental to Cavell's purpose. Strawson's view was an attempt to defend the concept of other minds as a primitive unity that describes both mine and the other's personhood. He believes we can have knowledge of other minds non-inferentially, directly, through criteria. The asymmetry between "I" and "(s)he" is surmounted because I must master criteria for the application of concepts to learn to self-ascribe thoughts, feelings and sensations. But given the publicity and primitiveness of these criteria, hence also of the concept of mind, I learn at the same time to ascribe to another (third personally) the same kind of inner life. In this case, from the fact that we ascribe a concept

¹⁴ In Strawson's words "it is perhaps better to say: X's depression *is* something, one and the same thing, which is felt, but not observed, by X, and observed, but not felt, by others than X. (Of course, what can be observed can also be faked or disguised.)" (p.109)

whose meaning is traduced as a primitive unity of mind and body, I know through criteria about other minds.

Strawson's strategy leaves us with almost all we need, but does not prevent us from asking questions about the nature of the knowledge acquired by means of knowing how to apply criteria. Criteria give us the grammar of a phenomenon, the conditions of application of a concept to it and its identity, but not *the existence* of what the concept is about. What should we do with the question of existence? Should we forget it? Should we try again to solve it? Is dissolving it a possibility?

Nagel finds in Sartre a turning tables attempt to answer the question of existence, while Cavell looks for an analysis of skepticism that dissolves the question of existence.

5. Nagel's Sartre and Look of the Other

In his book of essays *Secular Philosophy and the Religious Temperament*, Thomas Nagel honors and endorses Sartre's attempt to answer the problem of other minds. Sartre's *Being and Nothingness* 1943 is book full of rich insights on the nature of intersubjectivity, and Nagel acknowledges the way the French philosopher poses the question of otherness as remarkably original .

Sartre is aware of the difficulties of answering the epistemological problem of other minds. He does acknowledge the difficulty that arguments by analogy have in leaping from the other's behaviour to the other's mind, so he is also committed to a different strategy of answer.

Sartre's alternative is to look for "an unmediated, ground floor presentation of the Other, no less immediate than my understanding of my own subjectivity" (p.164). The idea of immediate awareness of the other calls for what Sartre dubs "the *Cogito of the Other*". The original, Cartesian Cogito in Sartre's opinion "is not a proof of my existence, but a way of seeing I have always known my own existence and cannot doubt it" (p.164), in other words, it is a sort of immediate recognition of my own existence enacted in the very act of recognition. The Cogito of the Other is no less than that, but it also involves an equivalent form of recognition that the other has been always living within in my own self. As a self-conscious being, I'm also capable of adopting the position of the other towards myself. I can be an object to myself, but even more, I can be a strange object to myself. This capacity can stay unarticulated, and that's why my knowledge of the other stands in need of enactment. Nagel calls this form of implicitness of the other a "pre-ontological" form of existence. I know it is not clear yet, but it will be.

Sartre's proposal is different from the ones we have seen so far. We have seen that the appeal to the publicity of the concept of mind was the strategy for those who don't want to fail by adopting the first person standpoint as the ground for answering it. Thus, appealing to the publicness of the concept of mind was a way to claim the equiprimordiality of the "I" and the

"other" - for I cannot understand myself without the other and vice-versa. Sartre's answer differs a little in comparison to the "public-concept" answer, specially when my understanding of the other's mind is constituted by means of public criteria.

He does not propose to find subjectivity, his own or that of the Other, unproblematically present as part of the natural order that we can all observe around us. That would be to ignore the problem by interpreting the self as a special type of object—in fact, to construe it as an example of being-in-itself. (Nagel, 2009, p.165)

The kind of answer given by Sartre is non-reductionist in a special sense. Although it looks similar to Strawson's solution, here in Sartre's conception the other is caught within the structure of my consciousness. In Strawson's argument, shared criteria give me the other as a "part of the natural order", as a different person that I can't fail to recognize as I learn to recognize myself. Sartre's "cogito of the other" must enact another consciousness within the structure of my self-consciousness and not as a different person who is part of this natural order. So, he is looking for something similar to the enactment of a self-conscious Other within me, a species of alienated self-consciousness. However we know that the other cannot be a homunculus that is surveilling my own consciousness.

Thus we get the question of how exactly "the Cogito of the Other" should work. Sartre's well-known example to illustrate this point goes as follows: Imagine I am looking through a keyhole from an empty passageway into an occupied room. Imagine inside this room a real spectacle is taking place. Something I am not allowed to know, but I have waited to see happen. I'm so absorbed that I can't be minimally aware of myself. I am completely engaged in unreflective consciousness of what I see and hear. Then imagine that all of a sudden I hear footsteps in the hall. Someone is looking at me! I am immediately covered by shame. I only feel the judgment of someone weighing over my head. So what this means? It means that I was immediately affected by the presence of someone, by some subjectivity, not by something that cannot judge me, but by something that has an appreciation of me.

Instead of a "I think, therefore I am", Sartre's Other looks at me and makes me aware of his look. My awareness of the "Look of the Other" is the awareness of a self-conscious being aware of its objecthood in someone's attention focus. As Nagel says this has a double advantage:

First, it provides an immediate and non-inferential grasp of the Other, and second, it reveals me to myself as an object in the world, and

not just a subject. The Look is the immediate experience of being perceived, and it gives me both the Other and my own objecthood all at once. In Sartre's terms, it reveals the Other not as an in-itself or as an analogue to the for-itself, but in the immediate being-for-others that is a further aspect of our consciousness and that is understood through an expanded use of the *cogito*. (idem. p.166)

In Sartre's understanding the recognition of the other in its immediateness shows that my understanding of the other is inseparable from the "structure of my own consciousness", therefore, it is as fundamental as the *cogito* itself. Nonetheless, Sartre's answer should not mislead us. This immediate awareness of the other is significant to understand the *idea* of the other and leaves the epistemological problem of other minds - the knowledge of the actuality of others - still open. The apprehension of other minds will only come by the actual experience of the Look of the Other, for only this gives me the awareness of myself as being the object of someone's understanding and the understanding that the other is a for-itself (a self-conscious) being. "The Look reveals me to myself as an object in the Other's world".

Although Sartre's answer gives us back the sense of intimacy with others that is often present in our relations of recognition, there are questions left open by Sartre's answer. Nagel himself asks "First, why doesn't the experience of the Look depend on a logically prior grasp of the Other, rather than providing it without presuppositions? Second, the idea of other subjects has to extend far beyond the range of those humans for whom I can become an object in the way that can make me feel shame or even feel perceived, so how can the Look be what reveals the general rejection of solipsism?"

I ask you to stay still with these questions, for they are to be answered in what follows.

6. Cavell's Knowing and Acknowledging

It is hard to fairly measure the impact and depth of Cavell's paper *Knowing and Acknowledging*. Centrally, his arguments deal with the -- up to that point, -- underestimated difference between knowing and acknowledging in analytic philosophy, claiming that difference to map into the difference between theoretical and the practical spheres of life. In the introduction to this dissertation I claimed to be interested in investigating the connections between epistemology and ethics; hence, if Cavell's proposal is sound, I will at last be able to start delineating that approximation. In doing that I intend to show how the ideas presented in *Knowing and Acknowledging* can reorganize and reinterpret the problem of other minds.

In order to understand the depth of Cavell's argument I would like to return to Frege's ideas about objectivity and the first person pronoun. Frege will be particularly illuminating in our

attempt to understand the difference between knowing and acknowledging, because his analysis of first person and third person pronouns in sentences using psychological predicates will give us a good picture of what is private and what is public - and how we should behave regarding what is private and what is public.

Frege distinguishes between "three realms" in his paper called *Thought*.

So the result seems to be: thoughts are neither things in the external world nor ideas. A third realm must be recognized. Anything belonging to this realm has it in common with ideas that it cannot be perceived by the senses, but has it in common with things that it does not need an owner so as to belong to the contents of consciousness. Thus for example the thought we have expressed in the Pythagorean Theorem is timelessly true, true independently of whether anyone takes it to be true. It needs no owner. It is not true only from the time when it is discovered; just as a planet, even before anyone saw it, was in interaction with other planets. (Frege. 1997 p.336-7)

The first realm is where the physical objects are - this table, this computer, etc. This realm is objective in the sense that we all can perceive the same things from our own perspectives. The physical world is open to everybody's view.

The second realm is the private world of the psychological. Psychological phenomena are typically feelings and sensations, and their privacy originates from the fact that they need a bearer. The psychological is different from the logical, and Frege could not have insisted more in this difference. This realm is not objective, rather it is subjective, in the sense that the experience of its objects is up to each subject separately.

The third realm is the logical. This realm has thoughts as its objects. We should not confuse the occurrence of a thought in someone's mind with the thought itself. Thoughts don't need a bearer, in Frege's understanding. That's why they can be shared by two or more people, whereas feelings and sensations cannot be shared.

But, if this division is sound, how should we understand the use of psychological predicates such as our thoughts about pain? What is the difference between self-ascription and third-person ascription of psychological states?

When I say "I am in pain", my thought *that I am in pain* is not, in a certain sense, about a public object of the world, my pain is only observable to others through my body, however my sensation is known by me immediately. This knowledge of my own sensations make it seem that I can know whether my statement is true or not, but this verification is not open to others,

since I can be pretending. But how can the truth of a statement be only accessible to me? Frege's early treatment of statements like this can be seen in "Logic", from 1897.

A sentence like 'I am cold' may seem to be a counter-example to our thesis that a thought is independent of the person thinking it, in so far as it can be true for one person and false for another, and thus not true in itself. The reason for this is that the sentence expresses a different thought in the mouth of one person from what it expresses in the mouth of another. In this case the mere words do not contain the entire sense: we have in addition to take into account who utters it... The word 'I' simply designates a different person in the mouths of different people. It is not necessary that the person who feels cold should himself give utterance to the thought that he feels cold. Another person can do this by using a name to designate the one who feels cold. (FREGE, 1897. p. 235)

When *someone* says "I'm cold" this does not mean that I can understand it on the same basis of *me* saying "I'm cold", for as Frege says in "On Sense and Reference" "sentences that have different cognitive values, have different sense." (Frege, 1892). So, If what you come to know by my saying "I'm cold" is different from what I come to know when you say "I'm cold", then those two sentences have different senses. The different cognitive value is attested by the fact that when I say "I'm cold" the person who is referred by the first person pronoun is Fernando and when it's you who say the same sentence, the person who is referred by the sentence is you, different from Fernando. Hence, these two sentences refer to two different objects and have two different senses. And of course, the "I" of your thought is not the same "I" of my thought when I claim to feel cold, for they are not coreferential expressions.

Consequently, if we are to preserve the truth or falsehood of I-thoughts, we must also admit the truth and falsehood of propositions such as

- A. I know that I am in pain
- B. I know you are in pain
- C. I know that s/he is in pain¹⁵

Therefore, at least *prima facie*, it seems that if I can know I'm in pain and you can know you're in pain, then it must be possible to know about the existence of your mind.

Cavell presents in *Knowing and Acknowledging* three positions concerning the problem of other minds. These positions result from the analysis of the three statements above. Not all sentences here will express knowledge in the same sense, even though they seem to at first sight. Cavell believes the skeptical and the anti-skeptical approach will result from the same

¹⁵ I will refer to them from now on as A, B and C.

understanding of what it is to know another mind, hence if they are wrong, and they must be, they commit a very similar mistake. A third option might be considered and this is going to be the apple of Cavell's (and my) eyes. I will consider each one at a time to then proceed to an analysis.

The Skeptic

The commonly known argument the skeptic gives us goes like this: (1) I know that I am in pain by being in pain. My introspection or self-consciousness informs directly about my inner states. In this sense, what I have is certainty in this matter. Also, (2) I cannot have the pain of others, but (3) all I have access to is the outward behavior caused by the pain. (3.1) I have to infer from outer to inner then, but (4) such behavior can also occur when the subject is not in pain. Therefore (...) The inference is not valid. I can't have knowledge of the other's pain.

The skeptic's argument is pretty convincing. It is a sound argument and we cannot battle the skeptic on his own field, Cavell would say. If we are to dismiss the skeptic's argument, then the strategy might be to show what goes wrong in his framing of the problem. But first, what is the skeptic's strategy?

[The Skeptic] begins with a full appreciation of the decisively significant facts that I may be suffering when no one else is, and that no one (else) may know (or care?); and that others may be suffering and I not know, which is equally appalling. (Cavell, 1969. p.247)

The skeptic's strategy calls our attention to the possibility of never noticing that we in fact don't know anything about the other's pain, so he rejects the possibility of knowing that the other has pain (like in sentence C above) and takes as a paradigm of knowing what it is to know about someone's pain to be that I, in first person perspective, truly know I have pain (like sentence A above). In a nutshell, the skeptic's requirement takes the sense of knowing someone's pain to be the same as *my* knowledge that I'm in pain. In what follows, notice how he dismisses all possible answers to his requirement.

Among the oldest strategies against skepticism one could put the skeptic in contradiction with himself by alerting him that it doesn't make any sense to argue with an automaton or with a brute creature, if I am one. However, the skeptic's ready-made answer to this: "You understood the problem. It's not that I know you're not a human. I just can't be sure about what you are. So, I'll react to the appearances as if they were true, even though I can't have *certainty* about them" The problem with this objection against the skeptic is that he is ready to accept that we live in an illusory world, so you're not shocking him, for he is precisely claiming that reality is shocking. His

solution will be that "for the sake of convenience" it does not make sense to take his argument as a lifelong torment, we should take it as a realization of our lack of sureness about the true nature of reality.

Another strategy against skepticism consist in perceiving that we commonly refer to the same pain in ordinary language. Since we perfectly agree in ordinary language, it is meaningless to put doubt on what we really want to say. Hence, if I have a headache and I describe it to my friend as a pain in the frontal face muscle, above my right eye, then it seems that he could react saying "Man, you won't believe, but I'm feeling the exact same pain". Of course, my friend is simply saying that he feels headache, the same kind of pain. However, in the eyes of the skeptic this is not certain. If I have my pain and you have your pain, how can we know we feel *the same pain*? So, the sense in which I can have your pain is a limited one. Pains and material objects are different when it comes to the epistemological access we have to them, hence different criteria of identity apply in their cases. We cannot have quantitatively the same pain - one and the same, as we can have the same car if we share it, for my pain is felt in my body and your pain in yours. We can only qualitatively have the same pain, e.g., pain in the elbow, mine in my elbow, yours in your elbow. That does not solve our problem, for I still don't know with the same certainty if you really have the pains I have. The skeptic is still smiling, since he is certain that we won't find anything so undoubtful as the certainty one has about one's own inner experience.

But if we are always falling short of certainty, then we should step back and notice that this is precisely what is defeating all our arguments. The requirement of certainty is precisely what cuts off our arguments in their general attempt to solve the problem. It encapsulates me to my experience, given that I can never be sure about the other's mental life. Hence, it seems necessary to analyse the relation between certainty and first person statements.

"The form of the words 'I know I am in pain' is senseless, as an expression of certainty", a contender might say. The idea behind this is that there are statements that do not make sense, because there are truths that it does not make sense to tell myself. This conception of meaning is tributary of a certain pragmatic model of meaning, where meaning is not only thought in terms of syntactical conformity, but also in terms of language use, in content vehiculation, tone and color. According to this sort of criticism, telling me, for the sake of my own knowledge, that "I know I am in pain" is not only non-informative, but also an empty linguistic action/performance. The skeptic can argue back, saying that its emptiness does not entail its falsehood. This strategy is known as the anti-skeptic way out of the problem of other minds and the fundamental claim is that the problem is senseless, because it is erected on a senseless conception of certainty. It will be interesting to see in more detail in what it consists.

The Anti-Skeptical Argument.

One of the possible readings of the problem of other minds consists in the rejection of the skeptical first-personal paradigm for pain (A). The anti-skeptical will argue that our true paradigm of knowledge of pain is the third personal (C), for first personal statements about the knowledge of my own pain (A) are nonsensical. But, what is the structure of an argument that rejects the skeptical argument?

So, the anti-skeptical argument starts stating that (1*) it only makes sense to speak of *knowing* where there is room for the contrast to *merely thinking that one knows*. But there is no room for such a contrast in case of one's own pain, for knowing that I am in pain equals having pain. (2*) I can have the same pain as another person, for qualitative identity is the only valid sense of "same" that can work here. Numerical identity does not make sense. (3*) There are criteria for applying the concept of pain to another person by observing them. If this person holds her cheek I can know that she has toothache or pain in her mouth. (4*) That I may be wrong in some cases (as in pretense), does not show that I can never know. Therefore (...), skepticism about other minds is meaningless. There's no puzzle about the other's pain.

The anti-skeptic claims a sort of problem solution, but the implication of solving the problem in this way is that we now seem unable to make sense of the (prima facie) *obvious* fact that I know when I'm feeling pain, even though I don't utter a sentence to that effect. The proposition "I know I have pain" isn't just a bunch of meaningless signs, even though it might not serve to tell me that I know I am in pain, perhaps it can serve to make a point to some very weird person who doubts that I can know that. But coincidentally this person is the anti-skeptic who's trying to convince me that I can't confer meaning to these words.

Now I would like to start the analysis of both the skeptic and the anti-skeptic's problems. We can notice that when the skeptic and the anti-skeptic argue, they both share the use of the phrase "can have pain" in their (respective) second premises, but they use it differently, with different meanings. In the first case the demand is for a numerical identity, in the second case the demand is for a qualitative identity. So, which one is the "correct" meaning?

Cavell will say: none. Both the skeptical and the anti-skeptical second premises will lead us to understand the expression of the other's body in different and incongruent perspectives. As Cavell puts it:

"At some stage the skeptic is going to be impressed by the fact that my knowledge of others depends upon their expressing themselves, in word and conduct. That is surely an essential fact to be impressed by. And then he realizes that the other may not in fact express himself, or that his

expression may be falsified (deliberately or in some other way); and that again is undeniable. (idem, p.254)

So, the bodily expression for the the skeptic means the impossibility of knowledge, for even though the body expresses true pain, the body can always fake pain, hence only from the inside one can truly know whether one is feeling pain. If we project the same worry for the anti-skeptic, the bodily expression of pain represents the possibility of knowledge, since it is only there that the requirement of third person perspective (C) fits. However, for the anti-skeptic first-personal perspective on my own pain is not an object of knowledge. Cavell will then argue that if we came to such an incongruence of perspectives the reason for their opposition might be one floor below the surface where the conflict is observable.

The skeptic's mistake

First we should notice the skeptic's mistake: a correct observation is interpreted as an expression of an epistemological limit.

(...) I am filled with this feeling of our separateness, let us say -- and I want you to have it too. So I give voice to it. And then my powerlessness presents itself as ignorance -- a metaphysical finitude as an intellectual lack. (idem. 263)

The skeptic mistakenly takes our separateness as an intellectual lack, as an intellectual incapacity to know the other's pain. But our separateness is a condition of our individuality, not a lack. Perhaps the experience of separateness comes in the form of intellectual lack, since the temptation of knowing the other in the intimacy of his thoughts seems to result in frustration and awareness of finitude. Our powerlessness to have each other's pains is not a lack in me, nor in you, it is just something we are not entitled to know. The lack of direct knowledge of the other's inner states brings a consequence: the skeptic stops treating behavior as expressive of pain. He separates expressive behavior and pain as two different possibly related phenomena and "scoops" the mind out of the body:

(...) behavior is expressive of mind; and this is not something we know, but a way we treat "behavior." The skeptic in effect goes on to say that we have no reason to treat behavior in this way. And is that false? -But what he turns out to mean is that behavior is one thing, the experience which "causes" or is "associated" with it is something else. That is, he stops treating behavior as expressive of mind, scoops mind out of it. My point,

however, is not to trace out the full extent of the skeptic's motivations; it is merely to deny that they, and what they lead him to, are senseless; or rather, to show that what he wants to know - namely, what it is we go on in the idea that behavior is expressive-is the right thing to want to know (idem. p.262)

When the skeptic stops treating behavior as expressive of mind, he is motivated by the fact that behavior does not *always* expresses what one feels. But is he justified in such generalization? Yes, behaviour can be misleading, but this does not mean that all pretence isn't a form of expression. Fake pain behaviour is a form of expression to our souls. To see this clearly we should notice that we don't say that the body itself is pretending to feel pain, nor that the mind is pretending to feel pain, precisely because none of them is autonomous to pretend. What feels pain or what pretends to feel pain is not the body or the mind, but a person — echoing Strawson. In the same way, a person is not behaviour or a soul, but an embodied soul. -- Remember when Wittgenstein asks to paint where is the pain in the body. Some people paint the brain, some the painful member, some other the connection of both. So, when the skeptic stops treating behavior as expressive of mind, we could say he is presupposing something less than a whole human. He picks one aspect of an inherently doubled-aspect being, as a person, and separates what is a unity.

This unity is the unity of a myth. Cavell says in *The Claim of Reason*, body and mind (soul) are part of a myth that helps humans understand themselves (CoR. p.364). Myths are narratives that explain by means of symbols the origin, development, and fate of things like psychological constellations, the history of a community, the origins of the rule of law. We commonly make appeal to myths when we have to explain those things, but we hardly can bring them to a clearer understanding. Mind and body are concepts that traduce our inner conflict between the identification to a natural order and the identification to the realm of intelligent creatures detached from merely causal relations. What philosophy did with the mind and body problem was the sin of substantializing characters of a myth and treating them as entities that we can have true knowledge. The skeptic's mistake is also to treat pretence as a lack of expressiveness, since he believes that soul and body can desynchronize. Actually, if pretence is an open possibility, this means that we can also express fake sensation. We can synchronize mind and body in many different ways.

Back to our skeptic problem, in the skeptic's point of view we cannot know whether something is veridical pain or pretence. Here we can remember McDowell's disjunctivism. For him, dream and reality are differently grounded experiences, and the disjunctivist argument is precisely to resist the skeptic claim that both partake from the same grounds. In the case of

skepticism about other minds, the body that pretend to be in pain and the body in real pain share the same external movements, analogously to the case of perceptual experience. Nonetheless, the body in pain, really feels pain; whilst the body that pretends, really pretends. But again the skeptic questions: How can I know which is which? Since I don't know, what should we do? Should we risk be the stupid that always believe the other is in pain? Should we as a way to prevent being evil always have the opinion the other feels pain?

Wittgenstein says that we are not of the opinion that the other has a soul. (PI.Part II, iv, §22) Neither I *have the opinion* that he is in pain, when I see him in pain. Our attitude towards someone in pain is not to formulate an opinion, but to help this person, to feel sorry, to feel empathy. If one fail to have one of these attitudes, what affects the person is something different. Wittgenstein calls this a case of blindness to the other's soul. (PI. §258)

Failing to notice someone's pains, feelings or mental life as trustworthy is a sort of blindness. Not a perceptual blindness, but more like an avoidance of recognition of the humanity of the other. It's not that one ought to feel always empathy for whoever cries, contorts or wince, but to intellectualize the pain of others as a first reaction reveals a distance to the ordinary human reactions. So, one does not see the humanity of others, because one is also afar from one's own humanity. Imagine what it is to live in a world where one is constantly thinking trying to discover if the other is truly feeling pain, because she can be faking pain! This world is theoretically possible, but ethically costful.

The cost of a world like the skeptic's is the blindness from one's own humanity. Wittgenstein explores this concept of blindness in connection to what it is to perceive a physiognomy. A physiognomy is what we perceive when we continuously connect aspects of perceived objects. Gestalt figures like the Necker cube or the Duck-rabbit can be seen in at least two ways, depending on how we connect their aspects. The different connections of aspects dawns on us the physiognomy of a different object. What these figures can teach us, Wittgenstein believes, is that humans inner and outer worlds are also aspects that work in connection. To perceive someone as having an inner life connected to the bodily outer world is to perceive this someone as a person. See things connected calls again for disjunctivism, for even in fake pain, the inner and outer are connected but in a deceiving way. In faked pain, the alleged connection between mind and soul dismantled by the skeptic isn't lost. So, it seems that we should understand deceiving pain not simply as a lack in humanity, as a failure in authenticity, or as reasons to be skeptical, but simply as another possibility for intelligent beings like us. Human life has contexts where all these possibilities are meaningful and there's no need to appeal to a common factor to understand them all. They have independent unity and separate existence. What the skeptic does with the possibility of pretense is not only to extract a

logical consequence from the possibility of always wronging, but he also breaks up his connection to the wholeness of human possible legitimate expression.

Another moral we have to extract about the possibility of pretence is that it shows not only that behavior is partly caused by the mind, but more fundamentally that the mind *has* a body, it owns a means of expression. Bodily ownership is connected to the agential capacity, which in its turn is the capacity for self-determination. Self-determination is expressed in action, which includes the capacity to determine one's own body so as to express one's mind. Therefore, when Cavell says that we *treat* behavior as expressive of mind, I read the word "treat" as a Wittgensteinian remark about our form of life. We treat body in the way we constituted its significance for us, in a way that expresses the agreement with our rules, institutions, needs and history. Since we are embodied beings having a body and making sense of it is part of what is to makes sense of ourselves as whole.

The Anti-Skeptic Mistake

As we have seen the dispute between the skeptic and the anti-skeptic is a dispute about what is the correct paradigm to understand psychological phenomena like pain. The skeptic takes the first person standpoint about mental occurrences construed as a privileged, immediate sort of access, as the model to understand the pain of others. The anti-skeptic takes the pain of others as the only possible model for understanding pain, claiming to be senseless to speak about *knowledge of my own* pain.

Cavell's goal is to show that both skeptic and anti-skeptic frame the problem of other minds from a shared principle. Both take the problem to be a problem of certainty:

(...) in fighting the skeptic too close in, as it were, the anti-skeptic takes over--or encourages--the major condition of the skeptic's argument, viz., that the problem of knowledge about other minds is the problem of certainty. At the same time, he neglects the fundamental insight of the skeptic by trying single-mindedly to prove its non-existence -- the insight, as I wish to put it, that certainty is not enough. What I mean can perhaps be brought out this way: In concentrating on the skeptic's apparently impossible demands (and neglecting what may be the insight which produces those demands) the anti-skeptic concentrates on the first-person half of the problem of other minds, to the neglect of the third person, as though half believing the skeptic's repudiation of the third person. (idem. p.258)

The anti-skeptic fights his opponent from too close, so to speak, and that makes him blind to the fact that they share assumptions about the requirement to answer the problem of other minds. The anti-skeptic must acknowledge that he cannot be sure about the existence of other minds, that's why he accepts that his knowledge is precarious in nature. The skeptic has a discovery in hands and his discovery is based on the impossibility of having certainty about other's mental states, at least as the certainty I have in a first personal perspective about my own mental life. The anti-skeptic combats the first person model of knowledge, but unnoticedly shares the basic assumptions, comparing first person to third person.

To show how this requirement is problematic, Cavell asks us to imagine a thought experiment. Imagine two brothers, First and Second, and imagine that a strange phenomenon affects their lives. First has a normal body and feels pain whenever it is inflicted in his body. Second, on the other hand, is not normal. He feels pains only when his brother is feeling pain. Actually, he feels exactly the same pain his brother feels, but if someone hits Second's body, he does not feel anything. Whip Second himself and he won't feel it. So Second only feels pain because First feels it.

Cavell's example shows us that in such an imagined situation we could have pain in *this body and in that body*, but we no longer would have *two owners of pain*, since we can't any longer tell whether Second would say "I'm in pain" or "My brother is in pain". Hence, we cannot expect Second to stick with only one of those forms of expression, for he does not necessarily have to adopt a third or a first person standpoint about his own pains. What is significant about this double possibility is that he can say both interchangeably, which entails that our normal concepts would not apply, for the distinction between first person and third person would collapse.

Now, it seems that in such a situation the skeptic would achieve what he wants. Second has First's pains, so he *knows* the ongoings of his brother's minds. Surely it cannot be denied that what the skeptic wants is precisely the capacity to penetrate the other's intimacy. Nonetheless, is that an accurate way to understand the situation? Cavell will say that we should analyse in more detail the grammar of pain ownership in this case. First has a normal body, and even if he knows that his brother feels numerically the same pain he feels, First will still take the pains he feels as his, despite the fact that his brother feels them too. First knows his brother's pains just as we know, that is to say, third personally, and only because Second expresses it in his behaviour. The fact that they occur infallibly synchronized does not affect the fact that First's knowledge is intellectual and distanced in nature.

Switching to Second's perspective does not help the skeptic as well. Here again Second knows First has his own pains, but what is baffling is that Second's pain no longer contrasts with First's pain. The conditions under which we could contrast one pain and the other are simply

missing. What is the sense of "my", in Second's mouth? When he says "my pains" he is never talking about the pain inflicted in his body, but only felt in it. So, the double sense of "my" that unites the inner and the outer is missed in Second's experience of pain. Assuming that I am Second, I no longer have criteria to establish the difference between my pain and his pain. Second's knowledge of First's pain is "too immediate" and First's knowledge of Second's is "too intellectual", Cavell will say.

The teaching we can draw from this mental experiment is summarized by Cavell as a reminder of our individuality and personhood.

But how shall we understand this wish for a response to my expressions (of pain, of any region of the mind)? Does it suggest that our concept of my knowledge of another is bound up with the concept of my freedom, an independence from the other, from all others-which I may or may not act upon? What is this "knowing a person"? (idem. p. 253)

What the skeptic and the Anti-Skeptic share is a sense of obliviousness, a lack of awareness of our individuality and that our mental life is the life of a person. The demands for knowledge that the skeptic presents are too strong and they do not make justice to the ways persons relate to each other when they are in pain. Knowing what a person is demands the awareness of the separateness and freedom that bounds each subjectivity.

Given the analysis of the Skeptic and Anti-Skeptic mistakes, we are in a position to understand what Cavell's alternative answer to the problem of the existence of other minds is.

Cavell's Alternative

The problem of knowing other minds is only possible if we forget that in a basic level our relation to other minds is not one of knowledge, but a relation of acknowledgement.

(...) there are special problems about our knowledge of another; exactly the problems the skeptic sees. And these problems can be said to invoke a special concept of knowledge, or region of the concept of knowledge, one which is not a function of certainty. This region has been pointed to in noticing that a first person acknowledgment of pain is not an expression of certainty but an expression of pain, that is, an exhibiting of the object of knowledge. There is an analogue to this shift in the case of third person utterances about pain. (idem.258-259)

Our knowledge of other minds is not a function of certainty, as modern skepticism has framed the issue. Certainty is precisely what the skeptic has asked for in his argument given what he understands about first person authority. His demand is so high precisely because the first person perspective on mental phenomena cannot be projected onto others, so she is led to scoop mind out of behavior when it comes to understand the third person attribution of mental states. In treating behavior as no longer expressive of a mind, he falls in doubt and never gets back into a state of certainty. Cavells points out here that what pain behavior shows to the person who feels it, is not certainty about pain, but pain itself.¹⁶

Cavell's criticism is based mainly on Wittgenstein's understanding of pain ascriptions. In PI 244, Wittgenstein describes the teaching of pain-behavior as a modification of the instinctive, natural expression of pain. What is first expressed by crying and contortion, say, is then modified into a linguistic expression like "I'm in pain" or "I feel pain". This modification changes our capacity to even talk about what was the bruteness of the feeling before acculturation. All we can say is that education can change raw nature into the conventions of our form of life. Obviously that causes the sensation that nature is lost, but the challenge here is precisely to see that nature is again in front of one's eyes. A highly sophisticated expression of pain may even change what pain is for oneself, whilst a completely suppressed expression of pain may turn pain into an almost unexisting sensation -- as in the case of the super Spartans that don't learn or completely suppress the feeling of pain¹⁷. The difference is not only between an expression that is not linguistically articulated and a linguistically articulated expression, but perhaps between ways of feeling determined by precise contexts.

Given the transformations that culture cause to the expressions of pain, we can't forget that both the cry and the utterance of pain are still expressions of pain, for both are still forms of address and avowals of pain that I publicly express. The publicness of my expression constitutes a form of address. Not the kind of address where one is trying to share thoughts structured linguistically, but a sort of engagement in second person relations. If I say "I'm in pain" to a friend, I'm not just trying to show him how great is my awareness of my feelings and emotions, what I'm doing is actually asking him to *do* something if my pain is caused by him, or to *do* something that is not under my power to do. If he can't do anything about my pain, at least I ask for his empathy. As Cavell says, my expression of pain makes a claim upon him and ask him to respond -- hopefully by taking care of me, but also his answer could be something closer

¹⁶ Here I think a connection with Interaction Theory (IT) is possible, since as we said earlier, this theory claims direct knowledge of other's sensations to be possible.

¹⁷ Putnam, Hilary. "Brains and behavior." *Readings in philosophy of psychology* 1 (1980): 24-36.

to "I don't give any importance". His knowing I am in pain demands a wider appreciation that does not stop in the mere cognitive act, but stretches to a demand for action.

Now notice that the problem of other minds has always taken the contrast between the first and the third person to be the cognitive bridge, for what one knows with certainty first personally is what one wants to know third personally. Cavell's intention is to deny this way of understanding the problem. He wants to show that if someone is a "s/he" for you, then at the moment this person says "I am in pain" s/he addresses you and treats you as a second person - as a "You". This claim made upon me demands acknowledgement of the pain of the other, be it a reaction of empathy, be it rashness and recklessness.

(...) your suffering makes a claim upon me. It is not enough that I know (am certain) that you suffer - I must do or reveal something (whatever can be done). In a word, I must acknowledge it, otherwise I do not know what "(your or his) being in pain" means. Is. (This is "acknowledging it to you." There is also something to be called "acknowledging it for you"; for example, I know you want it known, and that you are determined not to make it known, so I tell. Of course, I do not acknowledge it the way you do; I do not acknowledge it by expressing pain.) (idem. p.263)

Acknowledgement

Acknowledgement is something that we do. I may know that someone is suffering, but I can fail to acknowledge that, because I can stay cold and take no reactions on the basis of my knowledge. As Cavell's example clearly puts, when I know I am late, but I don't acknowledge that, I don't change my attitude in respect of my belatedness. But if I know I am late and I acknowledge this fact, I do something in this respect. Acknowledgment is something that provokes inner change, is something that we do to ourselves and to others.

Also, not all forms of acknowledgement demand sympathy. Your claim of suffering may go unanswered even though I acknowledge you are in pain. But suppose I get pleased by your suffering, since you lied to me and ruined my life. "The point, however, is that the concept of acknowledgment is evidenced equally by its failure as by its success. It is not a description of a given response but a category in terms of which a given response is evaluated." -- Cavell says. The difference between failing to know and failing to acknowledge is that "A "failure to know" might just mean a piece of ignorance, an absence of something, a blank. A "failure to acknowledge" is the presence of something, a confusion, an indifference, a callousness, an exhaustion, a coldness." (p.263-264)

The failure to acknowledge is a failure to exist in some aspects, it is the failure in bringing up possible relations towards oneself and towards others, since Cavell seems to understand the capacity to acknowledge someone as the capacity to *do* something to others and to oneself. If you don't do that; if you don't have this practice, you don't exist in this field of action. If you don't acknowledge your son, you fail to exist as a father. If you fail to acknowledge someone's feelings and emotions, you fail to exist as a moral being. We usually are naturally led to step into this kind of practice which is acknowledgement, but humans can withhold themselves from this and they can destroy their own capacity to acknowledge other by twisting their own capacity to express their thoughts, feelings and sensations.

Just as, to say that behavior is expressive is not to say that the man impaled upon his sensation must express it in his behavior; it is to say that in order not to express it he must suppress the behavior, or twist it. And if he twists it far or often enough, he may lose possession of the region of the mind which that behavior is expressing. (p.264)

I can become a stranger to myself, voluntarily or because life has led me to that. In the same pace I can become a stranger to others. I lose track of the humanity of others because I don't take neither me nor them anymore as expressive creatures. My inability to express myself reflects in my ability to takes others as expressive.

Part 2
Chapters 3 and 4

Chapter 3 - The Epistemology of Recognition

Spiritual emptiness is not a blank"
Stanley Cavell, *Knowing and Acknowledging*.

1. The Epistemology of Recognition

The first chapter of this dissertation was dedicated to investigate the problem of the *existence* of other minds. If other minds do exist, our philosophical wish is to *know* about their existence. However, after many attempts and proposed solutions to the problem, the analysis of its roots and motivations shows us that we cannot answer the problem in the absence of a deeper understanding of the philosophical confusions, demands and presuppositions that motivates the problem of other minds itself. Our conclusion in the former chapter was that knowing other minds is not a function of certainty - as Cavell states -, but a problem of acknowledging the other in a specific relation. Instead of a form of relation where I think *about* the other, acknowledgment is a form of thought *towards* the other.

The development of this conception of the problem of other minds also represents a shift concerning what is the object under investigation. Initially philosophers tried to reach a final and certain proof of the existence of other minds by means of arguments from analogy. This form of argument is a symptom of the modern quest for certainty mentioned above. The modern spirit we inherited led philosophers to treat the human body as a material object, similar to any other distant object of the world, leading them to forget the expressive nature of human behavior and, therefore, one of the most relevant features that constitute the concept of a person. The shift discussed in our former chapter implies that the cognitive relation with other minds becomes a function of the concept of acknowledgement, a different cognitive category that emerges when we step in *I-you* relations that entails a moral dimension.

As we have seen in the last chapter, Nagel's *Cogito of the Other* introjected the other into the structure of the I, making our awareness of the other constitutive of self-consciousness. In that sense no actual contact with others is needed in order to have awareness of other minds. The other is primarily *hypothetical*, for s/he resides within the structure of my mind. Therefore, my addressing the other is mediated by the primary hypothetical awareness of the other that I have within me. Cavell, on the other hand, claims that the nature of my relation to others is only

analyzable if I can be in the presence an *actual* other — if my claim upon the other is heard by someone. When the skeptic doubts the existence of outer world, it is part of his procedure to create a generalization from what he dubs as *best case of knowledge*. This is precisely the case when, for all in the world one is in front of an object, but the skeptic doubt still remain as a threat -- given that one always look at objects under a perspective. The external world skeptic's recital cannot be reproduced when it comes to the skepticism about other minds. The generalization that constructs the argument for external world, cannot be used for the case of other minds. Each person can be only known in its individuality, for there's no best case of knowledge for persons. The way a person manifests herself is always specific and unique. So, Cavell's idea is that the claim of the other upon me that leads me to an inner change is always dependent on the present, always in a specific moment of who I am and who the other is. The change it provokes in me is something that on the one hand has a passive dimension, for, as Cavell says, the expressiveness of the other makes a claim upon me, but on the other hand, my acknowledgement of the other is something active, something I *do* to myself and that changes my original position from an I-s/he to an I-you relation. In being addressed we step into a different relation.

This chapter will explore those questions in more detail. It will amount to an investigation about the epistemology of recognition. First I will follow two Hegelian thinkers, Honneth and Margalit, who show that recognition is a form of giving someone social validity. Social validity here has the same meaning that we have seen in the first chapter -- agreement and community -- however, Hegelian dialectics shows us the triangular structure of this stepping into communal agreement. The main question I will pursue here is *what is recognition?* My answer follows Cavell in *What's the Scandal of Skepticism?* regarding Wittgenstein's *Philosophical Investigations*. He shows there that recognition is a matter of accommodation. One must accommodate to the demands one does to oneself and to the legitimate demands of others, but in order to achieve this status the practice of address is an essential part of this.

We will also see in detail why address is the specific attitude that enacts our rational nature. And if this is so, then the next questions we will need to answer are "*why?*" and "*how?*". The nature of address will be further analyzed when we investigate Reid's "social acts of mind". He claims that a social act of mind is a different kind of thought that is irreducible to private acts of mind. On this matter Christopher Peacocke claims, in section 4, that the second person perspective, expressed by the second person pronoun, is reducible to privately thinking by means of first person pronoun and demonstratives. Therefore, he denies that social acts of mind, like address, are necessary to enact our rational nature. On the flip side, Sebastian Rödl

will argue, in section 5, that through joint activity we become aware of our own selves and of others. He claims that joint activity depends on sharing the same formal constituents in thought. The plural first person pronoun manifests the capacity for recognition in joint activity contexts. As we will see in section 7, Haddock doesn't disagree that address, thinking towards another, is the paradigm of recognition, however he holds that address can be present not only in intentional contexts, but as a non-intentional activity. We can think towards another non-intentionally because the human body expresses an act of address in its very nature.

The last part of this chapter, sections 8 and 9, will be an attempt to wrap up what in the end of the first chapter remained open. Given Haddock's insightful understanding of address, I believe we can also see some aspects of Cavell's understanding of skepticism in rich connections to one of the recent debates of analytic philosophy.

The Senses of Recognition

The verb "to recognize" has many senses¹⁸. This is observable when we analyse the logical form of this verb (x recognizes y) and notice that it brings different epistemic and ethical meanings when we substitute the free variables with different pronouns. A first and more common sense is recognition as the identification of an object previously known. This is a purely epistemological sense that involves memory and the capacity to correlate two separate experiences of the same object throughout time (e.g., I recognize this chair). A second sense of recognition occurs when one recognizes one's own mistakes - this is the sense of admitting something to oneself, acknowledging one's thoughts and mental events (e.g., I recognize I haven't been a good person). The third is a sense in which recognizing someone is to acknowledge the status and honour the other in his/her condition. In this later sense, which interests us the most, recognizing is not only having knowledge that something is the case - as a sort of knowledge obtained by means of observation and perception -, but involves a form of awareness that I and someone else stepped in relation (e.g., I recognize you). A further debate we will address in this chapter discusses whether the third sense of the word "recognition" also articulates the previous two other senses of the word. If recognizing someone articulates all the previous meanings, then we should ask *how* that happens. Is there any sense more fundamental to the others? Are they equiprimordial? Different constellations of how these three senses of recognition can blend together will give us different options to examine.

A first articulation of these three senses of recognition is present in Hegel's *Phenomenology of Spirit (1807)*. We can't ignore the enormous progress that the Hegelian

¹⁸ Michael Inwood, 'Recognition', in his *A Hegel Dictionary* (Oxford: Blackwell, 1992). In Margalit, A. *Recognizing the Brother and the Other*.

tradition of thought has made in this field of investigation, and whoever expresses interest in this area will always find in the *Phenomenology of Spirit* a rich and intriguing exploration of the concept of recognition. So, I consider that investigating Hegel's account of the concept of recognition is an obligatory first step in any satisfying account of the concept.

Before we go to Hegel, I must warn Hegelians that my intentions are not to debate the concept from within the boundaries that the tradition acknowledges. I think that in the long history of contributions to Philosophy, we don't need to always get back to the roots wherefrom that knowledge was first envisioned, instead we can judge what has been done, and progress from the fruits of this common enterprise. For that reason, I'll take two of the recent, but also greatest, expressions of the Hegelian thought: two works from Axel Honneth and Avishai Margalit. They will offer us the first topic I want to address: the relation between perception and the *act* of recognition.

2. Honneth: Recognition as Social Affirmation

I claimed initially that to recognize someone is to articulate in a unity the three aspects of human capacities: a cognitive relation to someone that involves identification, a reflexive relation to one's own consciousness and practical knowledge of someone's humanity. Honneth in his essay *Invisibility: On the Epistemology of Recognition* (2001) tells us about the situation of Ralph Ellison, a fictional character that happens to be a black man who's systematically humiliated, not by means of swears, offensive behavior or personal attack, but by the constant attitude of being looked through. His existence is constantly ignored by others who insist in not acknowledging his presence and behaving as if he wasn't present. Honneth will argue that the invisibility that massacres Ralph Ellison is obviously not caused by the incapacity to perceive him, but because there's an active attitude towards him that expresses the denial of his existence.

Perceiving a person is in a sense a passive act. When I am in front of a person and I don't have any ophthalmological problems, I can perceive the person in front of me. It's not only my eyes that must be in good shape, for perception includes my capacity to identify what is being present to my eyes. A minimal conceptual work takes place when we see something, as I have argued in unison with McDowell in the previous chapter. Hence, by perceiving a person, I mean to minimally identify the physical body of this person as the body *of a person*. Looking through a person is the expression of another capacity. It means to show disregard to a person who is present by behaving towards this person as if s/he was not in the room or not in front of me. "In this sense" Honnet says, "'looking through' someone has a performative aspect because it demands gestures or ways of behaving that make clear that the other is not seen, not merely accidentally, but rather intentionally".(p.112)

It looks artificial to say of someone who looks through somebody else that s/he doesn't know that it is a person that s/he is looking through. Hence we can differentiate cognizing a person and recognizing a person: *cognizing* someone as a person is an individual act that demands roughly two conditions: a) physical perceivability and b) individuation: the capacity for individual identification according to individuating properties. *Recognizing*, on the other hand, is affirming the social validity of a person after perceiving and identifying this person as a person.

Recognition extends beyond the cognitive act of an individual. It is a public expression "with the aid of actions, gestures or facial expression" of the fact that the person is "noticed affirmatively in the manner appropriate to the relationship in question" (p.111). Recognition also demands a medium of expression - gestures, facial expressions, etc, are necessary for recognition to be understood by the other -, since bodily expression is the socially perceptible element in recognition. These expressions are taken as symbols of my openness to that person, taking her as expressive and able to address me. Hence, if looking through someone is the expression of not recognizing someone, what is characteristic of this kind of action is precisely the expression of the denial of social validity.

But if our characterization of recognition is accurate, and if our way of distinguishing it from the simple act of cognition is sound, how should we explain the jump we make from the individual to the social act? The point here is not to deny that cognition does have a social aspect, but the emphasis lies in the fact that, for Honneth, recognition is possible *only* as a social act.

Honneth himself stipulates two forms of accounts of the difference between cognition and recognition. His first answer is that recognition is the addition of features to the act of cognizing. Hence, to recognize would be to identify someone as a person plus the expression of this identification. As he puts

In view of what has just been said, it could appear as if the act of recognition is due to an adding together of two elements: cognitive identification and expression. A certain person is first of all cognized as an individual with particular properties in a particular situation, and, in a second step, this cognition is given public expression in that the existence of the person perceived is confirmed before the eyes of those present through actions, gestures or facial expressions (Honneth, 2011. p.111)

This is not the first time that we stop by an "additive" conception. The additive conception of other minds, analysed in the first chapter, was a first attempt to answer the skeptic problem of other minds, but now we are dealing with an additive account of recognition. Here I

think it is clarifying to see things as Honneth does. He says that "the facial expressions and gestures with whose help human beings demonstrate recognition in direct communication, (...) cannot simply serve to reinforce an act of identifying cognition." (p.123) Our expressions of recognition of the other are not only signaling that "we know the other is there". They represent the actual openness to the other, they are the sign that we are no longer restrained by an egocentric perspective, which means that the other matters, and we are ready to do justice to the presence of that person. But what is that? What is to do justice to the presence of someone? I think that being open to others means to put myself in a state where I'm open to be addressed, but also to address others. It is a state where I'm active in relation to others -- I'm responsive to someone's expression. Perhaps going a bit farther than we can right now, I would say that it is to put myself in a state where I'm able to appreciate the other as an intelligent being. Hence, if recognizing someone is not only adding expression to knowledge, that must mean that when it comes to recognition, we cannot take cognition and expression as separate elements. Recognizing someone is a certain form of perceiving that is essentially expressive: "a form of perception with strongly evaluative features" says Honneth (p.125). This is the *transformative* view on recognition. A good case to observe it is in the relation between baby and caregiver:

It is not the case that we must first acquire a cognition that permits us to perceive in our counterpart a small child in need of help before we can then apply the appropriate gestures of encouragement and of sympathy. Rather, we seem as a rule to react directly to the perception of the small child with expressive responses in which a fundamentally affirmative attitude is expressed. The difference here may also be formulated as follows: in the first case, only a kind of cognitive conviction is demonstrated; in the second case, a motivational readiness is signalled directly. (Honneth, 2011, p. 117)

Hence, if recognizing someone means having an evaluative perception of the other, then recognition operates in the very act of capturing the presence of the other. I immediately have the other presented to me as person worth having my social affirmation of her existence. In contrast to the immediateness of this apprehension of the other, we have the case that calls the skeptic's attention.

Wittgenstein asks us in the PI 420 to imagine children playing in the street. He asks us to say to ourselves that those kids are just automata. If these words don't become mere empty sounds, given the meaninglessness these words have in confronting the vivacity of these kids in

the actual context, then you might succeed in producing an uncanny feeling. The uncanniness produced is precisely the result of adopting the skeptic's perspective on the world -- the dehumanization of the other and consequently the solipsism. This case, among other things, leads us to consider what is our natural posture when facing the other. Are we naturally and spontaneously ready to acknowledge to such an extent that dehumanization is the destruction or the fall of a natural capacity? Or the other way round: am I naturally prone to acknowledge only myself as human, while other's humanity is the result of a continuous struggle to leave "one's bubble"?

Primacy and Respect

Honneth's model of recognition based on the relation between the caregiver and the baby entails an immediateness of the relation of recognition that has two aspects: a) it has either a natural or a cultural primacy and b) it leads immediately to paradoxical moral considerations. These two aspects are fundamentally connected to the questions of the nature of perception as much as to the question of social validity.

The trust in this model leads him to a controversial thesis about the primacy of recognition over cognition. He says that "at least genetically, recognizing precedes cognizing insofar as the infant infers from facial expressions the 'worthy' properties of persons before he is in position to grasp his environment in a disinterested way.", but this is neither empirically confirmed nor conceptually necessary. Empirically things are a bit more complicated. The phenomenon of joint attention shows that the worth of a person can be simultaneously grasped with a disinterested appreciation of objects of the world. The studies by Tomasello on joint attention¹⁹ confirm a hypothesis envisaged by Donald Davidson that a triangulation²⁰ between two subjects and an object is a minimal cognitive structure needed to grasp the other's subjectivity. Joint attention is the capacity that two people have of paying attention simultaneously to an object at the same time that they are aware of each other's attention to the same object. This capacity is manifested in humans very early in life and is probably the basis through which other's subjectivity is objectively first grasped in babies. If triangulation is needed in order to have a minimal awareness of others, then the primacy of recognition is relative only to phenomena like "looking through", but not in relation to cognition of objects.

The second polemic in Honneth's account of the concept of recognition emerges from the unilateral way he understands the idea of "worth".

¹⁹ TOMASELLO - Joint Attention as Social Cognition in *Joint attention: its origin and role in development*, 1995

²⁰ Davidson, Donald. *Rational Animals* (1982)

[I]f recognition in its elementary form represents an expressive gesture of affirmation, it follows, to begin with, that it also represents a meta-action: by making a gesture of recognition towards another person, we performatively make her aware that we see ourselves obligated to behave towards her in a certain kind of benevolent way. (idem, p.120)

If recognizing someone is to express a behavior that affirms the other socially, then I affirm the humanity of the other at the same time that I understand this person as belonging to society. Honneth will say that understanding others as humans is the equivalent of expressing respect for the condition of others as equals. Equality here means that I must infringe my self-love in order to properly appreciate the humanity in others.

The manner in which Kant [...] speaks of something infringing upon my 'self-love' makes clear that here it is not the subject herself that imposes a burden on herself; it seems rather that the act of 'respect' as such has the active power, with the result that the suppression of egocentric inclination in the subject takes place, as it were, necessarily. To this extent, it would also be a mistake to speak of a mere resolve to limit oneself, because in the expression of respect 'self-love' is already infringed upon. Simultaneously with the expression of respect, the subject acquires a motivation vis-a-vis the respected 'worth' to forgo all actions that would simply be the result of egocentric impulses. (idem, p.121)

The humanity of the other is the intelligible aspect that I can grasp of others. So, if I should treat others as rational creatures, and this means that I should treat others in the same way I treat myself, then the concept of recognition necessarily entails a symmetric relation - that the relation I have to myself of respect for the person I am is stretchable to others. Nonetheless, there's something lacking in this conception of recognition. The possibility of mistreating others and being evil would be meaningless if I didn't understand others as persons. If I intentionally wrong someone, and if I want this person to feel humiliated, this wish only makes sense if I attribute the capacity of feeling humiliated to that person - intelligence and evil can also work together. Thus, the paradox of recognition is that, if that concept is to remind us that others are humans, hence worthy of respect and seen as ends in themselves, why is it fundamental to recognize others as humans in order to be evil to them? If when I'm aware of the other as a human, I get immediately aware of this person as an end in herself, I immediately know I owe

respect to her. It seems that in order to be evil, I must conceal from myself that I know I ought to respect to her, but in order to do that I must also do something with my own dignity -- I must ignore things that I know. This moral paradox will be addressed in more detail in chapter 4, however I want this to be a first connection of theoretical and ethical matters.

3. Margalit and Asymmetric Recognition

One of the objectionable elements in Honneth's understanding of the concept of recognition is the primordially of the symmetry between the two poles of recognition. His account of recognition starts from the positive relation of mutual recognition and builds the concept of looking through as a further deconstruction of the previous relation. Margalit goes to the root of the matter by bringing back to the discussion Hegel's understanding of the relation of recognition. He tries to show that we can also begin with asymmetry and then build up symmetric relations.

Asymmetric relations are either reciprocal or nonreciprocal. Symmetry has to do with how we describe the positions of the relation. Father and son are in different (asymmetric) positions as to *recognizing each other as members of a family*, but both can step in this relation reciprocally. When a father recognizes his son as a member of his family and the son recognizes his father as a member of his family, both recognize each other but from different positions. Reciprocity has to do with obtaining conditions. A common reciprocal relation is being married to someone. It cannot be true that *x is married to y* and the same time *y is not married to x*. Marriage is necessarily reciprocal, but it is an interesting case, for it can be both symmetrical or asymmetrical, if we analyse it in terms of, for instance, genders (it can occur with two equals or two different in most western countries: as aa, bb, ab). Recognizing someone as a person can be nonreciprocal, but it necessarily has to be symmetric. If we are in a nonreciprocal relation of recognition, it can be either that I recognize you as a person and you don't recognize me, or you recognize me as a person and I don't recognize you. However, what we both recognize, even non-reciprocally, is the position of the other as necessarily being of a person in the relation -- there's always the recognition of at least a hypothetical symmetry. Recognition becomes reciprocal when it's mutual, but it not always obtains; when it occurs we can say that both parties recognize each other as persons, and also recognize themselves as persons.

Hegel's *Phenomenology of the Spirit* describes the struggle for recognition as initially asymmetric. An important step in the constitution of Hegel's notion of spirit takes place in the understanding of the Master-Slave dialectics, that operates in the constitution of self-consciousness. Self-consciousness, for Hegel, is not constituted in reflection about one's own thoughts (as it is for Kant), but in the reflection with other self-conscious beings. The reflection is

only possible because we are first presented to otherness as strangeness to oneself -- as difference. Thus, Hegel's dialectic starts from asymmetry. These asymmetric positions lead one of the poles to demand recognition yet unexpectedly, but still rationally, the struggle for recognition of difference ends up as the dissolution of difference and the constitution of sameness and identity. Now we will see in detail how this transformations take place.

I'll start with the Master's standpoint. About this Margalit says

While observing the slave working with the tool, what the master gains in understanding is the essential difference between the slave and the tool, even though he regards them both as instruments for gratifying his own desires. In handling the tool, the slave acts by practical reasoning. Acting by practical reasoning is one way of being free, while still being a slave. It is this aspect of the slave—the ability to act intentionally—that makes the master realize that he encounters a different point of view: how things look from the slave's angle that is different from his own. (Margalit. 2011. p.131)

So, the Master's first grasp of a difference between himself and another is already present in observing the slave in his handling the tool. The difference between the slave and the tool becomes the allegory of the slave's practical reason, for with it the slave demonstrates capacity, skill, and more fundamentally, agency. Of course, even though the slave is under the mastery of the master, he is understood as free in an intellectual sense - as at least capable of breaking away from the authority of the master.

Margalit will also point out that the sense of domination that the master has over the slave is different from the kind of domination that the slave has over objects, e.g., the hammer. The slave with the hammer shows the mastery of a technique, while the master's domination over the slave depends on the slave's recognition of the master's dominance (whether recognized as peaceful or violent). So, the master must demonstrate interest in the slave's subordination to him, he must demonstrate his wish to colonize the freedom of the slave. In this movement of expressing his wish for dominance the master becomes self-conscious of his needs, as he treats his slave as a means, the slave's resistance provokes in the master the awareness of his own desires; the master knows that he wants recognition from the slave's part. This discovery from the Master's part leads him to the understanding of his own beliefs, the belief that he wants his beliefs to override the slave's wish for freedom, but at the same time the Master becomes aware of the belief that only because the slave is a free intelligent being that can recognize himself under the rule of the master.

Therefore, the master's realization of his own self-consciousness undeniably needs the aid of the slave's self-conscious condition. However, once the slave can notice such an insidious desire expressed by the master, the slave is also able to process a change in his condition.

So, what is the slave's process of self-consciousness constitution? How does the slave acquire self-knowledge from his angle in the triangle? The slave initially has only his own capacity to work -- which translates into the capacity to obey the master's commands and handle the tool. As the slave acts with the help of the tool, the conscious and wishful life of the master is contrasted with the inanimate tool. If the tool is not properly handled it does not complain, but as the master judges the slave is not doing something the way he wants, he will express the feeling of being contradicted. Then the slave discovers, in his struggle to express his own desires, the desires of his master and he discovers them as of a different kind from his own desires. The master's desires are powerful, they override the slave's desires -- at this moment the slave knows that his life is ruled by the master's wishes. At this point the master's desires of dominance are contrasted with the slave's desires of staying alive. If we were to talk about slavery as it really existed we know that the master must have inevitably asked himself "How will I control the mass of slaves? I can't torture them too much, because this revolts them and damages my own means of production; but I must demonstrate that they are governed by me, so they must know the range of my power". The slave on his side asks "Can I get rid of slavery? Should I risk my life and try to kill my master or should I accept being subjugated?"

So, the odd situation here is that on the one side the slave is a rational being, so he is in a sense free, but he is treated as a means; he is governed. On the other side, the master is also a rational being, free and capable as an agent, but he needs the slave as a means -- what makes him dependent. As Margalit says "His [the slave's] are desires to survive: to do all he can and at all costs— including the heavy price of being subjugated to the will of the other—in order to stay alive. The master, in contrast, is willing to take tremendous risks just to be recognized by him (the slave) as his master." The initial asymmetry between master and slave becomes a symmetric wish to survive. They mutually depend on each other in order to survive, but we should not take the struggle to survive as purely organic and biological survival. The master-slave dialectic is also an allegory of psychological survival, for survival here means to be recognized. In both poles of the relation the self-conscious knowledge of their own existences is only enacted by the other's recognition, and the denial of the recognition of the other represents a "death of consciousness". Why? Because here recognition is precisely what mutually animates ("gives life to") the minds of the relation. The discovery of self-consciousness is a second birth and undeniably the most significant one in our form of life. In this sense, the knowledge that you are self-conscious is the knowledge the you are alive. It is the knowledge

that you can make sense of yourself. Losing the knowledge that you are alive is then possible, because if no one recognizes your existence, then your subjectivity is dead to others. You feel incapable of meaning anything to others, because nothing you do is meaningful in a public sense. So, if you're not taken as a rational and free being by no one, your rationality is concealed to others. This possibility is metaphorical, but no less dramatic because metaphorical, the possibility of knowing of your own death - you become an observer of your own death in life. But that's not the last consequence. The last step in the degeneration of your own capacity for self-consciousness is the concealment of your rationality to yourself - the death of your own subjectivity to yourself - the feeling of never being safe between words, realizing you don't know how to speak your mind. The lack of recognition of others can make your own capacity to express your mind to yourself only darkness. You feel you're no longer accessible to yourself.

Notice that throughout the process of recognition, what both poles of the relation get is full distinguishability from material inanimate objects: they are no longer identifiable with dead matter. So, both are led back to distinguish themselves from the tool that initially was the means of expression of practical reason by the slave. Both are to be treated as ends in themselves, differently from the hammer which is a sheer means to other ends. This structure is what I have called earlier the triangulation structure.

In what follows we will examine the structure of triangulation as a way to understand the logic of recognition. What we have seen with the help of Margalit's Hegel is that the dialectics of recognition involves shifts from asymmetry to symmetry and backwards. The switching positions involved in this dialectics reflects the three previously mentioned senses of recognition. Identification (first order beliefs), Self-Ascription of Personhood (second order beliefs) and Ascription of Personhood to Others (third order beliefs) are involved in Hegel's theory of recognition, since it reflects the structure of self-consciousness and the recognition of the other's self-consciousness. Axel Honneth and Avishai Margalit give us some of the essential features of the relation of recognition, since their works shed light on the dialectic of power involved in the relation of recognition and how *power* operates on the structure of desire and knowledge. What is still lacking, we could say, is a more detailed logical analysis of the thoughts that constitute the transaction between the two poles of the recognition relation. Now I would like to analyse what is the form of our awareness of each other in linguistic communication; which kinds of thoughts are necessary for the dialectics of recognition to take place.

Some philosophers will tend to think that addressing another amounts to the articulation of separate and individual acts of mind composed by first person thoughts and demonstratives - both self-ascribed and ascribed to the addressee. Others will understand that the second person perspective is the typical form of address, so it is a form of evoking the first person in

other self-conscious being, but not reducible to demonstratives - it is rather a form of doing together. Depending on how we understand this logic, the idea of social validation becomes more or less important. If recognition is possible with two subjects thinking separately, then recognition can happen separately and may not depend on any public act of validation. However, if recognition is only possible by means of social acts of mind and forms of doing together, then recognition is dependent on the fulfillment of a set of conditions. Address is our main subject from now on, since the property towardness of looking at someone is at the basis of recognition. Now we will see in detail the difference between different conceptions of acts of recognition.

4. Thomas Reid and the Idea of Social Acts of Mind

If recognition is, among other things, giving someone social validity, then you might agree that giving someone social validity privately is useless or ultimately an incoherent idea. We want the other to know that we have given him/her social validity. For this reason, a social act of mind is needed. Now I think it is fruitful to acknowledge that if what we can learn from Hegel's Master-Slave dialectics is the social character of recognition, then Thomas Reid's idea of social acts of mind, present in his *Essays on the Active Powers of Man (1788)*, is yet another good contribution to devote some attention to in this connection. The idea of social acts of mind comes from the contrast with solitary acts of mind. Solitary acts of mind can be performed in solitude, "without intercourse with any other intelligent being". (Reid, 330). In contrast, social acts of mind are those that *necessarily* imply social intercourse with some other intelligent being that participates in them.

As Reid says, among the solitary acts we can think of seeing, hearing, remembering, judging and reasoning. We can do them without the intervention of other people and we don't need others to fully accomplish them. On the other hand, asking for information, testifying a fact, giving commands to servants, making a promise or entering a contract, are all social acts of mind. These cannot exist without the intervention from other intelligent beings bearing a part on them.

Reid will also add further distinctions. He says that solitary acts don't need to be expressed in order to be meaningful. So, "expression by words, or any sensible sign, is accidental" (Reid, 330) I can judge, think, remember or hear without other people's intervention or knowledge of what I'm doing. Social operations of mind, on the other side have expression as an essential component. They "cannot exist without being expressed by words or signs, and known to the other party" (Reid, 331)

Therefore, Reid's *Social Act of Mind* have three requirements to be performed,

1. *Participation condition*: A social act of mind needs at least two participants.

2. *Expression condition*: A social act of mind cannot exist without being expressed
3. *Recognition condition*: A social act of mind cannot exist without being known by the other party.

Now we might ask: how can *others* know when *we* are performing a social act of mind? The second person pronoun is a grammatical marker of address found in almost all languages. This pronoun is used in the vocative, but also in declarative sentences in order to address others.

Reid's idea of social acts of mind now can be fully understood: once we address others and satisfy the three conditions presented above, we can see that social operations of mind are *irreducible and basic*. Irreducible because the whole set of their components cannot be reduced to only one and basic because in the absence of one of these components the result will be a different phenomenon.

Centuries after the Scottish Enlightenment the idea of social acts of mind gained importance once again. The progress in logic and the development of tools for conceptual analysis have given philosophy the possibility to refresh old subjects and reconsider insightful works under new light. Peacocke is one of the neo-Fregean philosophers that will bring back the possibility of social acts of mind into the debate. However, as we shall see in the next section, his Fregean commitments lead him to deny that social acts of mind are irreducible and basic. In his opinion, given that the "you" pronoun is analyzable in terms of first person pronoun and demonstratives, social acts of mind do not need to meet the 3 conditions mentioned by Reid to exist.

5. Peacocke: The logic of the Second Person

Imagine you are a soldier in wartimes. You are not in the battlefield, but your job is to take care of the headquarter walls and stay alert for enemies that may come across infringing the minimal space for the army operational base. You walk alone during the night doing patrol service. Most nights are quiet and silent. One day you suddenly hear a click, few feet away. You freeze. It is the click made by a rifle being switched from "safe" to "ready to fire". -- Along these lines Peacocke begins his illustration of what is to be aware of the other's awareness of oneself.

Peacocke will argue that there are different kinds of self-consciousness²¹, and one of them is called interpersonal self-consciousness, where two subjects are conscious of being present to each other's consciousness. He holds, differently from Reid, that interactions such as face-to-face conversations, email exchanges, eye contact with another driver as one looks to

²¹ His main examples are: Joint Attention, Linguistic Conversation, Ordinary Self-Consciousness and Mutual Concern.

decide who is to cross the street first and non-linguistic interactions between mother and a young child (among others) are phenomena reducible to I-thoughts and demonstrative thought.

Back to the case we began with, if the soldier patrolling the headquarter walls is right about being aimed, then he will step into a complex phenomenon of interpersonal self-consciousness after listening to that sound. But if he's wrong, all the intersubjective self-consciousness is nothing more than an illusion. A dog could have produced the sound by stepping on some dried tree branches while looking at him. In this case the soldier is aware of a subject of mental states, but not of a subject capable of first person propositional attitudes. The dog would be only conscious of a human, but the animal probably does not know that the soldier is a self-conscious being. Therefore, in the dog case, the soldier is aware of being perceived, but he cannot be aware of a subject that knows that he (the soldier) is aware of himself being perceived.

Given the dog example, we can understand why intersubjective self-consciousness is a symmetric phenomenon. Peacocke will argue that it takes two self-representing beings for intersubjective self-consciousness to take place -- hence, it takes at least two subjects representing each other as capable of having de se attitudes (first person thoughts is the most remarkable form of de se attitudes given our linguistic nature). But how is the process of self-consciousness ascription built in social interaction? We will set aside for a while our dramatic situation where the soldier listens to the rifle aimed at him; imagine a situation where a pedestrian (x) is trying to cross the street, but he sees a car slowly coming. He tries to see through the windshield to make eye contact with the driver (y). The first moment is only a moment of visual perception.

We can say that

- (1) x sees y
- (2) y sees x.

As Peacocke says, if both pedestrian and driver had only seen each other, this condition would be sufficiently weak so as to be consistent with each person thinking the other does not see himself. If we are seeking for confirmation about the awareness of y as representing the other (x) as self-conscious, this would be insufficient, given the pedestrian aim of trying to safely cross the street as well as for the driver, given his aim of trying not to get involved in any car accident. They both must make sure that they are seen by the other. Both will try to see the other seeing himself²². So, a first level of embedding will result in

²² The structure of the human eye is of fundamental importance at this point. As Kobayashi & Kohshima, (2001) noticed, the white sclera that is characteristic of human eye compared with other primates. This special feature makes human gaze potentially much more expressive and, consequently, a tool for social interaction. (Kleinke, 1986) The eyes have the unique capacity to be the organ of perception, but also of being a means of expression (Gobel, Kim, & Richardson, 2015).

- (3) x sees that y sees him
- (4) y sees that x sees him.

Now suppose this driver and the pedestrian are in a countryside road during the night. What could have happened, Peacocke will note, is that conditions (3) and (4) are entirely consistent with x and y both featuring merely as an indistinct object in the other's awareness. Put in another way, x doesn't know *how (the mode of presentation)* he features in y's consciousness and vice-versa. The pedestrian only sees the driver seeing him, but he does not know if he is seen *as a person*. Both x and y can figure as mere *material object* in the other's field of vision, and mistakenly not be taken as a conscious being. So, in the case of intersubjective self-consciousness it is important that both participants are aware of each other as a conscious subject. This additional feature leads us to the second level of embedding:

- (5) x is aware that (4)
- (6) y is aware that (3).

Since x's awareness of y seeing him is only possible if y counts as a subject in x's mind, this kind of intersubjective awareness can be attributed only to self-conscious individuals. As Peacocke says "When (5) holds, and you are x, you are aware that the other sees that you see him. What you are aware of is a state of affairs in which the other sees something to be the case that involves your being a subject (an involvement of which you can presume the other has some kind of appreciation)." (Peacocke, p. xxx)

Even though this level of embedding is a high level in terms of cognitive capacities, this is still consistent with many animal's kind of awareness²³. Conditions (5) and (6) do not involve ascription of self-consciousness to others yet, however they are the conditions for a leap in a self-conscious ascription to others. Peacocke will call ascriptive self-consciousness the sort of consciousness where the first person thought enters in the specification of how the other thinks of me thinking about myself. So, if "y is conscious of me as one who thinks of myself as "I", then y is ascribing self-consciousness to me.

Therefore, in ascriptive self-consciousness two further sorts of thoughts will show up. The first one involves the presence of the first person pronoun in the accusative form: "me"

- (7) That person sees me

²³ An animal that passes the mirror test exhibits a capacity to link information about a perceived individual to its own condition. But there are two ways of having information about one's own condition: consciousness-as-subject and self-consciousness proper.

A creature that lacks a first person concept can in principle learn to translate information about the properties of a perceived individual into consciousness-as-subject of the holding of those properties. There's no reason why this transition must be mediated by an identity judgment (colloquially: "That's me!"). So the fact that an animal passes the Mirror Test does not show that it possesses self-consciousness proper. What's required for this awaits further analysis.

In the same way, it follows from the second level of embedding that the other's awareness is that he enjoys a sort of seeing with a content something like (8), as thought by you:

(8) He sees me. (modifying the demonstrative into a pronoun)

Up to now, condition (8) only gives me the awareness that the other sees me as a self-conscious being. When I know I figure as a self-conscious being in the other's awareness of me, the other does not yet know if he figures in the same conditions in my consciousness. As Peacocke holds "The other is here ascribing states with first person content to you; but he is not yet attributing to you a mental state which involves ascription to him of use of the first person notion or concept." (p. xxx)

The final step in this two-pile embedding is only reached when the other sees himself seen by me as a self-conscious being. This is the equivalent to *mutually* ascriptive interpersonal self-consciousness, where a third level of embedding is fulfilled.

(9) You are aware that the other is aware that you're in a state in which you'd sincerely say 'He sees me'.

Notice that at this level the second person pronoun finally appears. However, its appearance is understood as the expression of a description like "this person who is addressed by my utterance and understands that she is addressed". Then, Peacocke's analysis of mutual ascriptive self-consciousness forgoes the use of the second person pronoun.

He agrees with Richard Heck's solution in *Do demonstratives have senses?* when he says that "'you' [...] acts as if it were a special kind of demonstrative, one that always refers to the addressee. So if you want an analysis of 'you', try 'that person to whom I am speaking'". Peacocke's suggestion is slightly different, for he notices that in Heck's analysis the addressee can fail to know that he is addressed. Therefore, the suggestion is that the word "you", when present in address, is only a social transformation of the thought "This person who knows s/he is my addressee" or in the point of view of the addressee "I'm this person who is being addressed".

But you might still be wondering: How can I fail to know I am someone's addressee? And why would this entail that the second person pronoun is reducible to the first person pronoun? The evidence Peacocke offers in support of his reducibility thesis comes mainly from John Perry's well known supermarket example. As John Perry (1979) shows, I can use the second person pronoun without realizing I'm referring to myself. Perry's example is one in which "he is pushing a trolley in a supermarket, seeing someone in a mirror at an angle above pushing a trolley from which a stream of sugar is pouring on to the floor. Perry may think 'You are making a terrible mess'. He might even shout out loud 'You in the grocery aisle, you're making a terrible mess' without realizing he himself is the person addressed in the grocery aisle. But Perry does

not think "I am making a terrible mess", because he doesn't know that in this case *you = I*. Therefore, I can use the second person pronoun not knowing that my addressee is precisely me. Notice that it does not make sense to try to reduce the first person pronoun to the second person pronoun. That would be the equivalent to knowingly talk to someone else believing this person to be me. If you can use the second person pronoun without noticing that you can be talking about yourself, this means that it is reducible to first person pronoun and demonstratives.

This misidentification of the reference of the "you" pronoun works pretty much in the same way that I can misidentify the "I" in some kinds of sentences. In the *Blue Book* (p.66-68), Wittgenstein identifies two uses for the the first person pronoun. The first he dubs "use as a subject", the second he calls the "use as object". The first form is immune to misidentification, for I cannot think of another person as being the "I" in thoughts like, "I'm feeling pain". In this case, there's no sense in questioning myself whether I am sure it's me who feels pain. The second use can take others to be me, for there's no immunity to misidentification. A good and plausible example is: Imagine I survived a car accident. All I see is an arm completely contorted which I link to the pain I feel in my arm. In this case I could have said "My arm is broken" while still held in the mess of car metals after the collision. However, surprisingly after being dragged out of the car I realize it's not my arm that is actually broken. Therefore, no matter how schizophrenic this experience may sound, I can think of another person's body to be mine and I can mistakenly take myself (my reflex, my arm, etc) to be of another person. The corollary here is that If I can make first person pronoun use as object, I can mistakenly think of me as someone else - as in the supermarket case -- which may not prevent me from unknowingly addressing myself.

Now, suppose you say to me "You are F". Your intention is that I should thereby come to know that I am F. This is only possible if I know I am your addressee -- If I identify with the "you" in question. However, I could have not noticed that you are talking to me. I could have confused the "you" in your sentence with someone else close to me. In this case I would probably think "He thinks this person is F." Notice that by "this person" nothing less than "This self-conscious "I"" is entailed. The possibility of misidentification is, for Peacocke, what justifies that "you" is reducible to first person pronoun and third person pronoun, for every time I think "you", I am actually implying the capacity that someone has to think of *him* or *herself* as an "I".

Peacocke's reducibility thesis denies that social acts of mind are like Reid's threefold conditions establish. He holds intersubjective self-consciousness can be resolved into a constellation of individual and separate operations -- piles of thoughts that don't need to be recognized by the other to be meaningful. Perhaps the three conditions -- *participation, expression and recognition* -- are set aside because in the Fregean tradition the main concern of intersubjective dealings is to overcome the content privacy of I-thoughts. Nonetheless, the

separateness that each thinker can find him/herself in is a reason for skepticism, for we can always ask "was there recognition of the other's thoughts?".

Metaphysically we can say Peacocke's position has disadvantages, for success in communication seems essential when we are trying to account for intersubjectivity. It seems to me that when we try to give an account to others or to ourselves of what we have been doing while in a conversation with friends, when we marry, etc, it is part of the constitution of *meaning of our actions* that the other collaborates with her gestures, assent and attention. We usually don't rely merely in the possibility of uttering our thoughts in these social situations -- nobody keeps telling stories, nobody considers him/herself married, if the other person just runs away. It is not the same action even though the shared thoughts are formally identical. Even though sharing thoughts is a legitimate worry that is central for the possibility of communication, it seems irreproachable that if address is to be a meaningful activity, then, as Mathias Haase says in *The Puzzle About Recognition*, it must be recognized -- like a gift: unless it's taken, it is not given (p.117).

6. Rödl: Recognition as Joint Activity

Rödl's main criticism of Peacocke's view is that a social act of mind like having a conversation, caring about someone, getting married, etc., is not something that can be resolved into two piles of separate embedded thoughts that are supposed to connect thinkers by the simple metaphysical possibility of thinking the same thoughts and attributing to others self-consciousness. Rödl holds that recognition is something that depends on spontaneity and joint activity. This conclusion follows from his account of how pronouns do their job in distinct sorts of address.

In *You and I*, Rödl starts with his disagreement with Heck, Peacocke's inspiration. Rödl says "'you' is not a form of 'this', but a form of 'I'. Or better: 'I' and 'you' are the same form and logical function" (p.1). This point can be illustrated in a speech acts like the following: I point to three different things successively, saying "This, this and this belongs to me". Nobody will conclude that those three demonstratives have essentially different meanings simply because they refer to different things. It seems reasonable to understand that they are the same form of ostensive thought, even though they may not refer to the same thing. "This" here is used to simply point to a manifold of material objects. Now if I say "You, you and I are going to the room next door,", I can't conclude immediately that the first "you" articulates a different form of thought from the second "you"; given that any of those people who listened to my order could have given exactly the same order to me expressing the same thought interchangeably. So, we should investigate if "you" expresses or not a different manner of thinking of someone from "I".

The "I" is not a "This"

Shortly reconstructed, Rödl's argument first distinguishes "I and This", then it identifies "I and you". The difference between 'I and This' can be grasped by recalling the classical case of Oedipus. The king of Thebes unknowingly refers to himself as "this person" when he banishes the murderer of Laius. So, the difference between "I and this" can be stated as: 'with "I" one speaks of oneself knowingly' (p.2). Oedipus thinks of himself as a "he" when he unknowingly sentences the murderer, for he is definitely not addressing the murderer. If he had known he was the murderer, he would have thought "I murdered Laius, therefore I must be banished". This means that when we think of a third person thinking of herself we can think this as either a self-aware practice or a self-unaware practice. Hence, when I say "he thinks of himself" in the case of Oedipus, he is not thinking of himself in the same way a thought using the first person pronoun would articulate, for he doesn't know he refers to himself. Rödl suggests that when we refer to the third person knowingly referring to herself we can use the form "He_i" with the pronoun by means of which the subject thinks of himself.

The second step in this distinction between "I and This" is a formal one. Rödl claims that whoever uses the first person pronoun knowingly, knows of herself not by accident, but in virtue of the form of the thought. What this means is something that can only be clarified if we consider forms of obtaining knowledge, for if the first person pronoun use is to express some form of necessary self-awareness, then we must explain what is essential to it so that it results in such necessity.

Rödl's example to clarify this comes from predication. Imagine that you know that *a* is *F*. Think also that your knowledge that *a* is *F* may or may not rest on knowing that *b* = *a*. If my knowledge that *a* is *F* rests on my previous knowledge that *b* is *F*, then my knowledge is identification-dependent; if my knowledge does not depend on such condition, then it is identification-free. The importance of this distinction lies in the fact that if I know something about *a* based on my capacity to judge and to identify it with *b*, then this would be a way of thinking about *a* when I *think of it* in the way that *b presents me a*.

Now, if we are to compare the demonstrative "this" with the pronoun "I", then we should understand how demonstrative thought functions. Rödl follows Gareth Evan's understanding of demonstrative thought. In Evans' *Varieties of Reference* demonstratives give us the object of knowledge in an identification-free way, so that when I say, e.g., "This is *F*", there's no room for an identity judgment that will mediate my knowledge of the object I'm thinking about demonstratively. Accordingly, If I'm perceptually presented to the object that I know is *F*, then "I know that something is *F* by perceiving it [...]. There's no space for an identity judgment such as 'what I perceive is the same as this'. Knowing something by perceiving it, I know it *as this*. Demonstrative reference is reference mediated by sense perception.(p.5)".

Now think of the first person pronoun. When I use the first person pronoun I know I'm talking about myself, but not because I'm perceptually presented to myself. Hence, whenever one thinks of oneself as an *I* ("oneself", in Rödl's terminology), this use of the first person pronoun gives the person identification-free knowledge that it is about herself; that she is talking about. So, if 'I' was a form of 'This', then that usage would require an additional judgment that identifies "I" with what the demonstrative is referring to. However, if it is by means of perception that I know I'm talking about myself knowingly, then that would be accidental knowledge. (If knowledge of oneself were accidental, some people could never come across his/her true I, so she never perceived it or had the opportunity to at least demonstratively think of herself as an I.) Therefore, it's false that I use the first person pronoun knowingly because I have some sensory knowledge enabled by demonstratives. As Rödl says, "'I' refers in such a way that, not by accident, but in virtue of the form of reference, she who is referring is the same as she to whom she refers." (p.5)

The idea that the form of reference present in I-thoughts is not demonstrative leads us to explore a different model of reference. Given that demonstrative thought gives us sensory knowledge, demonstrative thought essentially gives us receptive knowledge, for sensory knowledge is only possible by means of receptivity (even though it may require activity as rapid eye movement and active powers of imagination). Also, we must acknowledge that there are basically two possible general descriptions of how we interact with the world -- as Kant has pointed out, we either are in a receptive/passive or spontaneous/active relation to objects. If this disjunction exhausts the logical space, then first person knowledge must be a form of spontaneity and activity. As Rödl put "One knows of oneself, not by being affected -- by something other or by oneself as other --, but by being active" (p.6). More about this is yet to come.

"You" is a form of "I"

Rödl's strategy from this point on is to show that you can implicitly have both pronouns "you" and "I" as elements in an *oratio obliqua* where the pronoun in use is "s/he". His example is very helpful to understand his point.

It may happen that Livia tells Roquairol that he is a traitor, saying "Roquairol is a traitor", to Roquairol, thinking, in the dark of the night, that she is speaking to Leon. If we rule out this possibility and report a second person statement, "You are a traitor", we use a special reflexive pronoun. (Rödl. *You and I*. p.7)

Hence, as Livia talks about Roquairol to Roquairol unknowingly, she is addressing him unknowingly as well. Of course, if she had known he was in front of her, she would have said "You are a traitor", but since she does not know, we can say that she used the third person pronoun addressing the person who is identical to the person she refers to. So Roquairol is indirectly addressed as "you", under the third person pronoun "he". Rödl suggests that we must have a word for this case: "he_{you}". It is a "you" of an oratio obliqua.

We can complexify things a little bit. The special pronouns Rödl has created, He_{you} and He_I, can come together in a statement. It's not hard to imagine: consider the statement "Livia said to Roquairol that she loved him". Here the pronoun "she" is ambiguous. I tend to think that she said "I love you" -- what would have resulted in "Livia said to Roquairol she_I loves him". In this case she is not speaking of herself indirectly, but she is using the "I". This is opposed to, for example, an indirect reference to herself "The heiress of Willbury Castle (who, as it happens, is Livia) loves ...". There's possibly more ambiguity in the sentence "Livia said to Roquairol that she loved him". She could have said she loves him unknowingly. She could have said "I love the Count of Durberville" without realizing that she was speaking to him. But in the case Livia said to Roquairol that she loves him saying "I love you", then we should report the fact with the sentence "Livia said to Roquairol she_I loves him_{you}"

Before Rödl's final step, he needs to show that "he_I" and "he_{you}" are forms of "I" and "you". Once we understand why this is so, we can finally start to understand why, in his view, "I" and "you" are one and the same function.

Now, we are asked to imagine two characters named Paul and Paula deciding what to do in order to accomplish a shared aim. Their final decision is that he would do A and she would do B, but the pronouns here stated were left open as to how he and she think of themselves in this decision. Hence, Rödl suggests

It may have been like this. Paul said, "I will do A and you B. Agreed?" Paula answered, "Fine; you will do A and I B." We can use our special pronouns to signify that Paul referred to himself and to Paula in this way, and write: "Paul proposed that he_I would do A and she_{you} would do B." And we can represent Paula's answer, writing, "Paula agreed that he_{you} would do A and she_I would do B." (idem. p.8)

Rödl then finally shows us that when we think of Paul and Paula articulating their positions with respect of each other we actually understand that their thoughts share the same content. Their deliberation can be described as "They decided that he_{you} would do this and she_I

would do that" and "They decided that he_I would do this and she_{you} would do that". This shows that "he_I" and "he_{you}", "she_I" and "she_{you}" are one form.

Rödl's last step identifies "I" and "you" as the same logical function. In order to understand this, we should remember that recognition, if it is to be a social act of mind, must at least relate two people. The plural pronouns are the way in which indo-european languages articulate distinct forms of collectives in thought. Hence, going back to our first example, when I say "You, you and I are going to the room next door" this only means that "We are going to the next door". This is possible only because we can bear in mind a collective (or plural) first person thought (a manifold of singular first person) that shares with the singular first person thought the property of having the one who is thinking to be the one the thought is about, or in the plural, having the ones who are thinking as the ones the thought is about.

Hence, whenever we think "We are going to do this", our thoughts are, on my side "You and I are going to do this" and on your side "You and I are going to do this" -- formally the same thought. So, despite the crossed reference (for when I think "I" the referent is myself and when you think "I" the referent is you), when we think by means of plural first person pronoun we are having the same thought in what concerns its form. This is revealing, but why is it so obvious that the reference is crossed? How does each person become aware that she's not the "I" in the sentence when I utter, e.g., "I love you"?

Wittgenstein offers an insightful suggestion related to this point in a passage from the *Philosophical Investigations* in which he imagines a scenario with humans that do not have the practice of addressing others:

A human being can encourage himself, give himself orders, obey, blame and punish himself; he can ask himself a question and answer it. So one could imagine human beings who spoke only in monologue, who accompanied their activities by talking to themselves. -- An explorer who watched them and listened to their talk might succeed in translating their language into ours. (This would enable him to predict these people's actions correctly, for he also hears them making resolutions and decisions.)

But is it also conceivable that there be a language in which a person could write down or give voice to his inner experiences a his feelings, moods, and so on a for his own use? — Well, can't we do so in our ordinary language? -- But that is not what I mean. The words of this language are to refer to what only the speaker can know -- to his immediate private sensations. So another person cannot understand the language. (PI 243)

Wittgenstein asks us to imagine a community of humans that can only speak in monologue. In a form of life that is only capable of monologues you cannot address others directly, but it does not prevent you from thinking about others. However, let's try to imagine that these strange humans can only refer to themselves, as if they were imprisoned in I-thoughts. Now imagine you're one of them. So if you were one of them, you wouldn't know what it is to be addressed. So, whenever you hear a first person utterance, you take it as a thought of yours coming from the outer world. This is possibly contradictory, for the grammar of I-thoughts teaches us that the only possible author of these thoughts is you, but try to think that you were actually passively thinking about yourself. Imagine you have heard someone saying "I must go to the restroom" and you take it to be a thought about you with the weird sensation of not being the author of this thought. A good image of this comes later in the *Investigations* when Wittgenstein says "If I listened to the words of my mouth, I might say that someone else was speaking out of my mouth." (PI part II, x).

The strangeness of these images comes from the violation of the grammar of I-thoughts, but also from what it is to recognize someone (including myself). As Rödl shows, when we decide to do something we share thoughts with the same logical form, however we do this occupying different positions in jointly thinking these thoughts -- and that's why I don't take reference wrongly when you say "I love you". Also, if thinking I-thoughts is a sort of active and spontaneous form of thought, then talking to someone, addressing a person, is a form of "doing together" where we necessarily express and have to be recognized by the other in order the action to fully take place. As Rödl says

When I speak to you, and not merely unsuccessfully attempt to do so, then we are speaking to each other. That I am speaking to you and you understand me are sides of one act, a joint act of spontaneity: conversation, Gespräch. (Rödl, *You and I*. p.12)

Having words (or sounds) affecting my ear is not the same as *listening* to someone. If I *listen* to you in a conversation then I have knowledge about the conversation. If your words only *affect* me I'm not yet a participant on it. My listening to you is active listening in, e.g., a conversation, in listening to your saying "yes" when we get married, when I listen to your cry of pain. My being affected by your words is passive and not a form of "doing together". Being active in joint activity or being merely affected by others transforms the nature of the addressee from a distanced third person to a close second person. This is a criterion to understanding what makes the plural first person an actual unity.

The difference that this conception of recognition as the action thought by means of the plural first person makes can best be seen in the contrast it produces when compared again to Heck's definition of "you" as "the one to whom one is speaking". As Rödl says, "speaking with someone is not the same as being affected by his utterances. The former is necessarily an object of joint spontaneous knowledge, while the latter is not." Therefore, getting back to Heck, we cannot analyze "you" as simply "he to whom I am speaking", for this person may not be engaged in joint activity with me.

7. Clearing things up

Now I would like to compare Peacocke's and Rödl's theses regarding the way they diverge when it comes to the nature of intersubjectivity. Peacocke's theory of intersubjectivity is partially in agreement with Heck when he says that:

This description of what is involved in using and understanding the second person does not invoke a special second person concept or way of thinking. It uses only third person and first person singular concepts, and concepts of those concepts (and further thereof, up the Fregean hierarchy). (Peacocke, 2014, 245)

However, Peacocke thinks that we understand the second person as "this person who is addressed by my utterance and understands that she_i is addressed". Different from Heck, but in partial agreement with Rödl, this description uses both a demonstrative pronoun and a first person pronoun. The first person pronoun is an important addition to Heck's former idea, for he fails to recognize the need of the first person understanding of oneself as the addressee. Peacocke represents this by the formula

$$\text{You're } F = [\text{self}]_b^{\text{is } F}$$

He explains that in this formula "<...>" denotes the sense of "...", and "[self]" is the first person thought-type indexed with the object b. So, all b needs to do is somehow to relate to the sense of the first person pronoun in play in the utterer's sentence. Peacocke says that the utterer refers to the first person type of mode of presentation and by this the utterer intends that b is capable of employing $[\text{self}]_b$ in thought. This kind of neo-Fregean treatment of indexical content differs from Rödl's idea that "the capacity to know that thoughts stand in a certain relation depends on the capacity to share them" which can be understood as the *recognition condition* present in Reid's conception of social acts of mind. Peacocke's idea is that you can

know if my first person thought is of the same first person type as your first person thought without having the attitudes Rödls believe are necessary. The neo-Fregean philosopher endorses Vittorio Gallese's thesis²⁴ that we share a manifold of intersubjectivity (2005). This means that we, as some apes, share specifications of events described as "*x is F-ing*" and use this data, grasped from perception and intentional action, to create a sort of general perspective-independent map of space. This map of space, since it is perspective-independent, serves as the means for different subjects to regulate their own bodily image as immersed in an objective environment. As Gallese himself says 'Rather, the shared space instantiated by the mirror neurons blends the interacting individuals within a shared implicit semantic content' (2005. p.111). Hence, mirror neurons help in empathically project the other's body in the same perspective-independent space, but given that this projection is the projection of another being -- capable of intentions, desires, and so -- this space is shared as if having implicit semantic content, for each time I'm presented to objects I assume these objects have the same possible significances they have to me. Gallese's thesis agrees with Peacocke's understanding of intersubjective as separate perspectives referred by demonstratives, for in their hypothesis it is by means of the intersubjective manifold that we recognize that other human beings are similar to us making possible intersubjective communication, social interaction and mind reading.

So, the main difference between Peacocke and Rödl resides in the inclusion or exclusion of demonstratives as necessary constituents of the content of the second person pronoun. If demonstrative thought, as Rödl suggests, is essentially receptive and I-thoughts are essentially spontaneous and active, then address doesn't need any form of demonstratives, because address includes the other as an agent in joint activity in a way that s/he does not think her/himself as object of someone's thoughts. Peacocke thinks intersubjectivity in terms of separate thoughts and solitary acts of mind (to use Reid's terminology). So, no common act is needed in Peacocke's view. But if he is right in claiming that in address you have two piles of separate thoughts, how are we to understand successful communication in contrast with unsuccessful communication? It looks like whatever I think or do will always fall short reaching the other's mind.

Rödl's suggestion is more promising because it permits a better and broader understanding of both passive and active uses of the second person pronoun. His thesis focuses on address as the main and most distinctive form of active and spontaneous joint self-consciousness, which for him is nothing more than a form of practical joint rationality. This is not

²⁴ Peacocke position in this debate is tributary of Vittorio Gallese's 'the shared manifold of intersubjectivity' (2005: 111). Peacocke says he "[...]has just been supporting Gallese's position, for the special case of interpersonal self-consciousness, and in particular when it is attained by second order simulation."

to deny that I can address others as posited in the passive pole of the dialogical relation, for this possibility must also be open. However, in this case there's no need of recognition; in this case we can think of forms of second person pronoun use where no joint activity will result such as, e.g., "You have nutella in your cheek". If I address someone saying this, then I think of a person making her aware of herself as an "I" observed by me. Perhaps this first sentence opens a space for a conversation and a form of joint agency, but notice that according to Rödl's suggestion this is not from the start a form of joint action. The person will know that a state of affairs obtains by means of a testimony, which is a receptive form of knowledge.

The difference between active and passive address comes from the conception of address as an intentional action. As Rödl claims, when addressing someone both poles of the dialogical structure are aware of their positions in an action. So, participation of other also changes the way we think the prepositional arrangement of thoughts: When you thank someone *for* something, or advise someone *against* something, the action of address is not only directed *at* someone, but *to or towards* someone. Actions *towards* someone creates a metaphysically and linguistically distinctive entity; they put two or more people in joint activity. Directedness is what makes address an action that asks for a broader understanding that includes the social aspect and different criteria of success.

Finally, I would like to close this section calling attention to an aspect that Rödl does not explore in detail in his paper, but I see it as fundamental to understand the kind of criticism that is to come in the next sections. When Rödl takes address to be a form of joint action he's taking it to be an intentional action. His description of two subjects knowingly thinking themselves by means of the first person pronoun is precisely the picture of an intentional action. We don't take this into deep consideration when we try to understand the language game of referring to someone knowingly or not, but this can hide other possible conceptions of address -- and of what recognition is. Now, the kind of question that I'm left with is whether it is possible to think address as a non-joint action. Adrian Haddock argues in *On Address* that it is possible to understand other forms of directedness to others that include non-intentional actions. Now I would like to explain how I understand his view.

8. Haddock: Between Passivity and Activity

Now we want to understand Haddock's conception of address, which is derivative from his understanding of intentional²⁵ actions that do not reach full realization -- the action that for

²⁵ The subject of intentionality is certainly a very hard one, but we can work with a minimal definition of intentional action "as the realisation or execution of practical thought -- an intention, or volition or desire." (Haddock, p1). Traditionally, consciousness plays an important role in the description of intentional actions, for the capacity to relate an action to oneself as the author of the action or the capacity to place it in a justificatory set of premises where you recognize the volition to do x in your previous mental states are in general criteria to take an action as intentional.

some reason remained incomplete. This is important because if some acts of address are unsuccessful, then only the thought's pure structure is expressed, according to Haddock. If address's pure structure can be expressed both intentionally or unintentionally, then the means of expression of humans -- the body -- is itself an act of address, as Haddock claims.

Let's start from where he starts. Haddock relies on Hornsby's account of actions of address. She understand address as an action that meets Reid's three conditions: if a subject x addresses y , then x must have a volition to address y ; this action can fail, so it must be understood in terms of trying-to-do. If x succeeds it's because x has not only met certain criteria of what it is to address someone in a given context, but also because y has recognized that x is trying to address him/her. Then, if x succeeds in having y participating, expressing and recognizing, then x 's practical thought (intentional thought) was realized.

The successful case of address is what Rödl has analysed. His model is essentially intentional, but we can still ask what are the criteria of completeness and then this will open up to the possibility to consider what is a derivative case of unsuccessful and not fully realized mode of unexpressed and not recognized thought.

...this mode of intentionality must be exhibited not only in realised practical thoughts of the sort the account articulates, but in other contexts too; specifically, in unrealized practical thoughts -- pure intentions, or volitions, or desires. (Haddock. 2014. p.1)

But what causes an act of address to be incomplete? If it's not me who fainted or had a stroke in the middle of the process, I might have failed to single out my addressee. But why? Peacocke insisted that my address must relate to the *type* of first person thought, but as Haddock notes, on this level of generality intersubjective exchanges can not occur. A person may not relate to the sense of first person type evoked, because it lacks elements to be recognized by the addressee. A very common example of this difficulty can be seen when we talk to kids. A specific intonation, mode of putting words together in a sentence, a particular form of bowing is sometimes necessary to get the full attention of infants. In a Wittgensteinian jargon, we can say that a sort of blindness to some forms of expression can affect people so as to result in incomplete actions of address. So, the act of address, in order to be acknowledged must meet some conditions. I have to think and express my thoughts in a way my addressee can recognize himself in my act. For example, a thought that relates me to someone I'm trying to advise against the dangers of eating too much fat, or that relates me to the same person when I'm trying to declare my love, respectively represents the other in two different specific

ways. As Haddock says "It cannot be a merely general representation -- "I'm trying to address *another (indistinctly)*". I'm trying to address someone with whom I might or not have intimacy, someone of whom I can know the history, someone I perhaps haven't met in a year; and all this counts when I'm trying to touch a specific subject with this person. Therefore, as it seems, my address must be expressed in a very specific manner in order to be recognized. Recognition is only possible if my expression singles out the person that is my addressee, if my expression fits the kind of representation that the person has of herself or expects me to have of herself.

We can say, then, that address fails because one may not only make a demand for the type of first person way of thinking. The second person way is a way that demands some knowledge of *who you are*.

"I am trying to address you." This is not a perceptual-demonstrative way of thinking. And yet, just like such a manner of thinking, it is an element of singular thought. It is the first person way of thinking. What is distinctive about this practical thought is that it thinks in this way, not merely of the subject of the thought, but of another subject. So, there is something misleading about simply calling this the first person way of thinking. It is equally the second person way. (Idem. p.4)

There can still be space for *y* to wonder "Is he trying to address me?" -- *x* might not have noticed that he is still representing *y* in a way *y* herself doesn't recognize as accurate anymore. The blindness that can surreptitiously affect *y*'s self-recognition in an act of address toward her can have multiple reasons: *X*'s lack of knowledge, *y*'s lack of sympathy, etc. This entails not only that *x*'s volition must represent *y* in a certain manner, but that such a manner must be known by *y* in order to be acknowledged. So, in order for *x* to represent *y* in the correct way, *x* must have some knowledge *about y*, such that if *x* represents *y* wrongly, *y* won't be capable of identifying his/herself_i in the other's address -- *y* won't be capable of recognizing *x*'s address as *towards* her.

The fact that *x* not only thinks about *y*, but fundamentally *towards y* is of ultimate importance to what Haddock wants. As he says

Directedness to a fellow subject is a property of the practical thoughts that inform speech acts, whether or not these thoughts are realised in intentional actions. So, it cannot be that the present mode of intentionality is exemplified merely in realised practical thought; it must be exemplified in pure practical thought as well. (Idem. p.2)

So, Haddock believes thinking *towards* someone is a special way of thinking because it includes the other in your thoughts as another subject. Those thoughts do not need to be intentional, and they do not need to be formed in joint activity. The directedness to a fellow subject can be silent -- .g., my long days of work need not be acknowledged by my baby who is at home with his mom in order to be an act of address. My actions are towards him. Address must have directedness.

Obviously this poses a problem: As Haddock recognizes:

But this raises a question. It seems that x can have a pure practical thought directed towards y even if x does not engage in any overt addressive actions, or indeed in any overt behaviour at all which is in any sense informed by the thought at issue. And yet, it is a condition of x's having such a de re/de se thought about y that y is able to recognise that she is thought of by x in this way—whether the thought is realised in action, or expressed in behaviour, or not. How is this possible? For there to be something for y to recognise, x's thought must be in some sense publically available. (Idem. p.5.)

So, how can x's thought *still be available to y*, if address hasn't been acknowledged?

Haddock asks us to follow him in a very unexpected consideration. He says that "x has (or is) a body: x is present in the public world, and is capable of realizing thoughts in intentional action and expressing thoughts in behavior." He continues saying that "the difficulty is that it seems that x can have thoughts towards another that are neither realised nor so expressed -- and yet are such as to be apprehensible by the one to whom they are directed nonetheless." (§16)

Haddock claims that our status as embodied creatures expands the sense of address beyond being merely a character of intentional joint action where the other is explicitly articulated in the pronominal form. Address is to be found wherever thinking towards another is to be found. The father that works the whole day for his son, the teacher that worries about his students understanding and struggles to prepare a good course. All these actions are addressed, however not examples of joint action. Haddock understands these actions as expressives of address as well, just not expressive in the presence of the person to whom the act is aimed at. The body of whoever thinks of another person in its expressive nature "[...]is an act of address" (§17)

Haddock's thesis contains a degree of generality that can be dangerously taken for emptiness or vacuity, but I believe this is a risk that holistic accounts can take after doing the fine grained job. Since he does the finegrained job, we can understand wherefrom his conclusion comes. He is holding that we should not understand address as belonging either to the categories of passivity or activity -- he does not discourage what Rödl and Peacocke say, he seems to believe that intersubjective exchanges can occur either way, whether in joint activity or in a passive and receptive ascriptive self-consciousness. Instead, what he seems to be doing is avoiding the dispute about the true nature of address by calling our attention to a broader understanding of human expressiveness as such.

9. Humans Condemned to Expressiveness

Haddock's analysis of address is focused on cases where we fail to reach attention and recognition. The unsuccessful cases surprisingly led us to understand in more detail two features: i) that an act towards another has the pure form of directedness and ii) that failure in address comes from giving the wrong content to the act. Address must express a sense with which the other person can identify, since when we are addressed by others we are not only beings waiting for the type of "I-thought" sense to awaken us from the constant stream of our consciousness, but actually we are self-representing beings who in general expect others to follow this self-understanding as a means of expressing their own recognition of us.

Self-understanding here means that we are beings that forge narratives that try to account for who we are; but we have feelings that are triggered according to the history of our lives and thoughts about ourselves that we may not want to express -- the uniqueness that each person veils under the fabric of ordinary life.

The fact that we have to reach the other there where s/he finds him/herself represented in order to recognize and be recognized seems to force each act of recognition to penetrate a field of solitude and privacy. My uniqueness seems to be concealed to others in a first contact, hence it seems reasonable to say that it takes time to uncover the person I am. It is only by sharing things in everyday life that the specificity of my life will come to light. This does not mean that we can only address others when we know them well, but it means that there are specific ways we can (and cannot) talk to people unknown to us; because we can't talk about everything with everyone at any time.

But if time is essential for intimacy in human relations, it seems that there is a region of my life that I inevitably will/must conceal from others. I can avoid others, but I can prevent others to notice I'm avoiding them. If address can fail because I don't let myself being read by others, then we are back to the private language argument.

The private language argument hides issues concerning self-knowledge. Perhaps because it is too much on the nose, perhaps because the argument really shows up in Kripke's words with an appearance of a ("merely") logical problem. Cavell will be interested not only in the conceptual scheme that has to be recruited to "solve" the problem, but also in the specific aspect of the problem that allows us to understand things related to self-knowledge in the private language argument. According to Cavell a fantasy of denying the publicness of language moves us to endorse the claim that there are regions of our thoughts, language and expression whose meanings are not open to others.

[...] the wish to deny the publicness of language turns out, so far, to be a fantasy, or fear, either of inexpressiveness, one in which I am not merely unknown, but in which I am powerless to make myself known; or one in which what I express is beyond my control.

[...] A fantasy of necessary inexpressiveness [...] would relieve me of the responsibility for making myself known to others. (Cavell, 1979. p. 351)

If meanings were not public, but private, my authority over what I say would be limitless. And at the same time, if my authority over what I say goes to infinity, my misunderstandings with others potentially would go together. The problem with having complete authority over what I express is that this would entail the burden of never allowing others do the job of understanding me. The responsibility over meaning would be totally on the shoulders of each person that expresses something. In this sense, responsibility and authority would amount to be the same, as if no one were allowed to step autonomously in the practice of understanding someone else's words and thoughts. Every conversation would be equivalent to trying to penetrate in a mine field where whatever you understand can be denied. In such a skeptical scenario, is it possible to deal with such degree of responsibility? It seems that my responsibility over my expressiveness must be somehow limited in order to make sense, otherwise it can evaporate into thin air.

How can responsibility be limited? First we should distinguish responsibility and authority. My authority over what I express comes from what I consciously express and gloss as my expression. If I can identify myself as the author of something meaningful, I'm also responsible for it. Authority has here two meanings -- one is of allowance, permissibility and the other is one of authorship, being the creator. When one successfully projects (new) words into new contexts one is the author of a meaningful new step in language. It creates new meaning. Creation must be acknowledged, nonetheless. It must be traced back to some source of shared

rationality, otherwise the creation won't be understood, it will remain meaningless. However, not everything I am responsible for comes from my authorship. My responsibility spreads not only over what I am the author of, but over what I trust to the point of reproducing and sharing with others. What I perpetrate through tradition, institutions and habits is also (partially) under my responsibility. Hence, responsibility is wider than authority when it comes to consider my expressiveness. It seems, then, that my responsibility is shared with others.²⁶

But does responsibility over my expression mean control over it? Cavell conjugates responsibility and control. The fear he observes in admitting the publicness of meaning and language is the fear of losing things out of my control. Since I definitely express more than what I'm in control of, the publicness of meaning allows others to read my attitudes. My attitudes and my reactions were learned in my community, that's why others can read them. (as we have seen when we addressed the problem of the civilized cry of pain). This is contrary to the skeptic's wishes. He does not want to be responsible for what he is not in control of. He wants to acknowledge only the privacy of the content in order to have full control of what he means. This is what gives the skeptic so much authority, but so little power to communicate.

Losing balance of responsibilities messes things up. When we lose balance, our attunement is lost. We have to try to accommodate to each other again. If we don't succeed in accommodating to each other, then skepticism comes back. Skepticism, Cavell says, is the situation where accommodation is not possible. Address is constituted by this accommodation to the other, and in order to represent my fellows as s/he identifies her/himself, I must share a long list of things that Cavell more than once lists in his works: interests, judgments, impressions, needs, inclinations, desires, temptations, compulsions, surprises, moods tastes, curiosities, qualms, jokes, perceptions, games etc.²⁷ Acknowledge other humans in the precise position that they expect to be acknowledged is a job centered in the singularity of each person. As Cavell comments in *What's the scandal of Skepticism?* It can seem a miracle that on the whole we do not *have* to accommodate ourselves to one another in speaking, whilst at the same time there's no such thing as a *best case of knowing other minds*. There's no general

²⁶ So, If it is to be shared, it should equally reach some balance, for it is also in balancing participation in community that we can reach the meaning of what is to autonomously express oneself. Autonomy is desirable because it means that one is sufficient to the task - an autonomous person does not need another to be the means for expressing him/herself; autonomy demands that one doesn't take needlessly responsibility over others. My responsibility over others is expressed in terms of respect to the other's freedom. The balance that a community must reach is a sort of reliance that each person must have in the other's capacity to learn, to share and to make sense of herself. I'll talk more about this in the third chapter. Here I want to outline the future connections I think should be done.

²⁷ An emblematic moment where Cavell comments the multiplicity of possible aspects in which our form of life can disagree is in *The Availability of Wittgenstein's Later Philosophy*, page 52 of *Must We Mean What we Say*.

knowledge that can teach us how to deal with people according to each specificity. Each person is an emblem of her uniqueness.

So responsibility and publicness of language relate because the fear about the publicness is also the fear of being always somehow the source of my expressions, even of what I don't want to acknowledge as my expression. The wish located in skepticism is to endorse the idea that if my control over what I say is lost, then responsibility over my expressions should be lost too. In this case, the temptation is to refuse responsibility over my expression, to free myself of the weight of having to make myself known to others. To get rid of the possible significance that I have in participating on someone else's life. I'm incognito. I can't be spotted. Here I think we can conclude that if meaning is shared, then responsibility over it is shared as well. Meaning is the patrimony that we inherit, indeed it is one of the most important inheritances in our form of life, for it is the inheritance that puts each new individual in a position to self-constitute and constitute social bonds.

When Haddock says that address can fail, it is important to notice that it can fail because others sometimes fail to let me know them. A shadowy person projects the darkness of her own soul over her own body. The inexpressive body expresses a psychologically devastated person. Haddock's thesis that the human body is an act of address, should also enable the understanding of the lack of self-address -- the lack of inner dialogue. Not hearing my own inner voice and not taking responsibility over it makes harder for others to acknowledge me -- I can become a distorted expression of myself. This is not exactly an expression of something that I am not, but it is the expression of my effort to disappear. Hence, If the body is an act of address, it should also address others representing my lack of self-address. An inexpressive body is so dark that the skeptic doesn't see it anymore.

The Ethics of Recognition - Chapter 4

Giving a reason for something one did or said means showing a way which leads to this action. In some cases it means telling the way which one has gone oneself; in others it means describing a way which leads there and is in accordance with certain accepted rules.
Wittgenstein, *Blue Book*

1. Summary

This chapter will explore another aspect of the concept of recognition, not the theoretical, but the practical one. As I mentioned in the previous chapters, the bridge between the epistemology of other minds and the epistemology of other wills entails the continuity between philosophy of mind and ethics. In chapter one the theoretical question that led us to the philosophy of recognition was: how can I know if others have mental life? If they do, do they have inner experiences like mines? These questions, we have seen before, in the absence of a definitive answer, entail skepticism about other minds. The skeptic about other minds believes each person experiences one's own mind in absolute privacy. Since the alleged certainty with which I know my inner states cannot be replicated when it comes to the knowledge of others, the idea of knowing with certainty that other minds exist revealed itself beyond any possible proof. By the end of part one, what we grasp from Cavell's *Knowing and Acknowledging* is that the cognitive dimension of other minds is only one among other possible relations we can have with others. The analysis of ordinary language shows that we constituted a form of life where addressing other people is more an act guided by relations of recognition (acknowledgement in Cavell's jargon) than any other forms of relation (like the cognitive the skeptic puts pressure on). So, we were led to the concept of recognition as the concept that better grasp the dimension of our relatedness to other minds. Now we face what is an almost inseparable aspect of the same concept: the ethical question of recognition.

If recognition is not only an intellectual form of relation to others, but also it entails some practical knowledge about the way I ought to *treat (how one should act towards)* persons, then relations of recognition are not only the answer to an epistemological question about other minds. It opens up the ethical dimension of what is to recognize a person with different wishes, volitions and desires from mines. The concept of a person, in the theoretical dimension makes me aware of an intelligence, whilst in the ethical dimension, it makes me aware of duties, responsibilities, emotions and so on.

So when I acknowledge others and think them as persons, does this change the way I relate to them? What is the ethical normativity (duties and responsibilities) connected to the idea of another person? Throughout this chapter I hope to answer these questions.

First I will start the debate with the classic distinction between consequentialism and absolutism (also known as deontology). The way we will address these matter is by closely reading Nagel's *War and Massacre*. In this article we will focus on why recognition entails absolutism. Since absolutism points to the importance of knowing what is the dignity of a person, our interest throughout the chapter will be to investigate which are the sources of the normative claims recognition entails. The second step on this final chapter will be to contrast Darwall's arguments with Korsgaard's. Their opinions differ when it comes to understand what is the source of our capacity for second person standpoint - which is precisely the standpoint acknowledgment requires. Darwall believes we should think the categorical imperative counterfactually according to our idea of what a person is -- from this we can extract what is to recognize a person in her dignity. Korsgaard believes we learn to adopt a second person standpoint over ourselves through action. So, the roots recognition are in oneself's practice of learning how to be a person. By the end of this final chapter I tried to show that these two solutions are not in necessary opposition, but they point to different practices of recognition.

2. Dignity and Rationality

In *War and Massacre*, Thomas Nagel inquires the possibility of respecting human dignity in extreme conditions. His reflections about what is to be seen as an enemy in war contexts tries to differentiate persons *qua* persons from persons *qua* any other quality or feature one may have. His goal is to understand what can be taken as disrespect to human dignity in contexts where much of the worth of human life has been lost. From this inquiry, Nagel will claim the value of absolutism²⁸ from both agent-relative and agent-free perspective. In the world as we know it, violence and war are not always avoidable; hence, if your country goes into civil war, it's possible that you no longer will be in conditions to decide whether or not to kill a neighbour, a classmate or a co-worker. People you know can become true strangers depending on the side they take, the opinions they support. Killing someone can be the only way to survive or to grant the life of your beloved ones. What Nagel investigates is whether it is possible to keep a sense of respect for the humanity of others even in such a miserable and dreadful context.

The idea of respect for another human life and dignity is contrasted by Nagel with other form of moral thought in which doing what is more honorable is not the goal, but ideally the goal is to maximize the good. The classic locus of this debate revolves around the so-called "trolley cases", however Nagel delivers us that old dilemma (between consequentialism and deontology) in a brand new and more dramatical shape, where the subject of human dignity is in a first layer of analysis already lost, given the raw and cruel state of war.

²⁸ Absolutism is the term used by Nagel to make reference to what other authors call Deontology, Categoricalism or Ethics of Duty.

Consequentialism, also known as Utilitarianism, can be summarized by its lemma "Maximize utility". The principle of maximizing utility recommends that agents should promote the greatest possible good as the consequence of their actions. Consequentialism then, thinks action looking forward to what will happen in consequence of action. So, the ends have greater importance than any other aspect involved in the action. Also, given that the focus is on the consequences of actions, the evaluative standards in a consequentialist practical rationality assesses their total outcome. So, the agent's motivations are not under consideration, but only or centrally the overall good obtained in the resulting states of affairs.

Consequentialism entails no restriction on means to achieve the greater good. But means are available throughout time and in restricted quantity. So, the criteria of magnitude and time apply in consequentialist deliberation. According to consequentialism, you can choose action x instead of y, if x generates a less evil consequence than y. But also time is a criterion for deliberation, for if in the long term the consequences of y override the immediate evil, but x don't, then y is a better action than x according to utilitarian principles.

Although our moral intuitions frequently work out making use of consequentialist principles, when it comes to more specific contexts, they might not be enough. For example, these two principles were heavily applied by governments to justify torture, as in the case of President Bush's administration²⁹. Many people thought that muslims represented a threat to the life of innocents. So, they believed that torturing suspects could lead to a lesser evil, for, even though you can torture the wrong person and wrong her, to many people it could prevent a greater evil from happening. That is roughly a consequentialist reasoning where no restrictions on means apply in order to produce the (supposedly) greater good.

Absolutism does not deny consequentialism, but offers a different evaluative criteria for morality. It values not the state of affairs resulting from the action, but other set of components of action. What is central to an absolutist standpoint in morality is the evaluation of the agent, its motivations and the merit of his or her actions, hence good and coherent motivations lead to good actions. Differently from Consequentialism, in Absolutism not all means are valid in order to achieve *some* good, since there are actions that are valued as intrinsically offensive to morality. For example, in an absolutist point of view, torture and murder are absolutely prohibited no matter what the ends or consequences you are trying to achieve with those means. The reason is that no action is good if it violates human dignity and treats other humans as means. Human dignity is an absolute value that cannot be negotiated.

The Paradox of Deontology

²⁹ <https://www.theguardian.com/world/2009/apr/17/torture-methods-interrogation-george-bush-approved>

The assessment that the absolutist position tries to bring about the value of human life. You cannot disrespect the autonomy and the free will of others in order to get information for your own interest, the absolutist says. Doing this is the same as failing to recognize the other's humanity, the fact the other is not a means to anyone's ends. But we could ask ourselves in a real situation "why not murder this one in order to prevent 100 or more from being murdered?" (Isn't it just vanity to want to keep one's own hands clean? Or isn't it blind obedience to norms and absolutist values?) The absolutist position is paradoxical if it requires in an actual situation that one refrain from murdering one instead of saving many lives, for deciding not to kill *actually* entails losing many other lives. We tend to (reasonably) think that the effort to respect one (or a few) life (lives) cannot cost the life of hundreds.

Nagel calls it the Paradox of Deontology and his effort is to think that we are not always in a position where the principles of absolutism and consequentialism can be thought in simple and clear contexts. Sometimes we cannot avoid wronging someone and, as a fatal consequence, we should acknowledge that the world can be an evil place. Nagel believes that sometimes "the utilitarian considerations favoring violation are overpoweringly weighty and extremely certain – nevertheless, when that special condition is met, it may become impossible to adhere to the absolutist position." (p.126)

It is normal to feel a bit confused in such a scenario. So, if consequentialism allows one to kill in order to save lives, does that mean that the consequentialist does not see value in human life? Not exactly, this is not the right view about consequentialism. Actually, the value of human life is embedded in a relation to what is desirable as the best outcome. However, in consequentialism the value is *relative* to what we may call the maximizing the good. Absolutism thinks the value of human life as an *absolute* value, that can override any other principle. The problem here is that we don't want to deny the absolute value of human life, but at the same time the consequentialist seems to be in a better position because he actually can prevent the catastrophic death of many.

Nagel's answer tries to restate absolutism harmonizing it with consequentialism. He says that "Absolutism requires that we *avoid* murder at all costs, not that we *prevent* it at all costs." (p.132). So, avoiding the evil in a world that can be evil is to accept that we have to conform our actions instead of getting stuck in "frictionless" principles. We can't be afraid of action because we don't want to wrong others -- actually we should question ourselves if the best we can do includes wronging others, is that still to be thought as a bad action? Perhaps not. We must act on the basis of the best deliberation we have in hands. Of course, the point here also is to show that the world might be larger than my capacity to think it, I'm not always in a position where I can freely chose the consequences or the factors involved. So, given that evil is not always avoidable, sometimes I won't be able to prevent it at all costs. But isn't this exactly

failing to recognize the other's humanity as a surrogate principle? Is recognition something one can manage without when it comes to debate moral matters?

Instead of adopting a purely consequentialist resolution that would allow torturing others, Nagel suggests that we should understand the problem not in relation to which facts will be the best outcome of our actions, but instead regarding these actions as relations to persons. In other words, we should manifest an attitude toward *the other person involved* rather than just to the *situation* alone. What does that mean?

Nagel will try to show that even in adverse conditions, one can acknowledge the other's humanity. He gives us two examples of how we can treat enemies during war. One of them treats enemies as humans, the other as real threats.

So we must distinguish combatants from noncombatants on the basis of their immediate threat or harmfulness.

Children are not combatants even though they may join the armed forces if they are allowed to grow up. Women are not combatants just because they bear children or offer comfort to the soldiers. More problematic are the supporting personnel, whether in or out of uniform, from drivers of munitions trucks and army cooks to civilian munitions workers and farmers. (Nagel, 1972, p.140)

Many people work for the combatants during the war, but not all those who work for the enemy soldiers should be considered as soldiers. On the one hand, there are people who work for the soldiers providing food, delivering letters from their beloved ones and helping the enemy soldiers *as persons simply*, who need basic stuff in order to exist. On the other hand, people who transport military supplies, guns, bombs, drive soldiers to the battlefield, even though they are not in the frontlines, they are cooperating with the soldiers not only as humans, but fundamentally as military -- so as contributing with possible killers. So, the first point to consider is how the population is involved in wartimes. Some people cooperate only in the most general dimension of life with soldiers - the basic needs of humans --, others contribute with the specific task of military activities -- the basic needs of soldiers.

This distinction is fundamental to understand that the life of enemy soldiers is to be considered under a variety of aspects. In wartimes rumors, passions and fear complicate a lot our understanding of what's going on and what one ought to do. It can be the case that a General think the only solution for him to advance his troops is to bomb certain areas where not only military personnel lives, so killing innocents is in this case necessary. So, supposing he

knows that enemy troops are hidden there, he might think he have to make it, even if he does so trying to avoid killing innocent people. These justifications are what he calls administrative, for "not everything that happens to others as a result of what one does is something that one has done to them" (p130). Here the death of innocent people is a means to achieve an end. It does not mean that it isn't wrong to kill them, but it means that one can be in the situation where it is necessary to act in consequentialist way and still consider issues concerning human dignity.

Even if you recognize that doing something bad or painful is necessary, you can still show some appraisal for the other's life and dignity. Imagine that you are in the middle of a fire situation in a sinking ship. If you abandon a person in these situations saying "I have to leave you in order to save the kids", this person should understand your choice. Similarly, if one subjects unwillingly a child to a painful surgical procedure, one can say to him, "If you could understand, you would realize that I am doing this to help you." One could even say, as one bayonets an enemy soldier, "Sorry. It's either you or me." However one cannot say to a prisoner that it is absolutely necessary that you torture him in order to get some names of suspicious people. I think what Nagel has in mind is that we can differentiate even when it is hard to see the difference. *Ought* not always *implies can*, Nagel says (p.144). A consequence of this will be a double conception of agency and obligation: agent-relative and agent-independent. The agent-independent obligations are the ones we have to other agents abstracted the possible contradictions that reality might engender. These obligations are the oughts we not always can accomplish, for they belong to us as subjects in the view from nowhere. On the other hand, the agent-relative obligations are the ones we experience given our the specificity of each concrete life. These are the oughts we can accomplish. The absolutist position must be understood taking into consideration these two point of view.

Nagel believes that even though it does not get us very far, the practice of justifying an action to someone that can truly understand you is a different kind of practice than that of having to administer the death of many. Justifying your actions to the world at large, as in consequentialism, is different from thinking your justifications in terms of address to the person you are wronging. The first form of justification treats humans indistinctly and from afar, the second treats a person in the proximities. This difference is fundamental to our goals.

We will see throughout the chapter that the act of address makes us aware of a normativity that is only accessible due to our second personal relations. The practice of justifying your actions to others in an act of address is what Nagel is trying to understand as the possibility of recognition even in morally wrong actions. When you think of the other person as a rational being, you access the nature of the other and his proper dignity, even though reality does not always permit both to coexist without wronging each other. The investigation of how address is supposed to elucidate our moral nature is what I want to understand in what follows.

The first topic to be examined is Stephen Darwall's understanding of morality as an exercise of empathy and imagination. The second is Christine Korsgaard theory of interaction with oneself as a way to learn to interact with others. We have seen in chapter three that address is the way we can recognize and be recognized by others. Nonetheless, recognition has left behind some clues about the ethical importance of this concept that now I explore.

3. Darwall and Second Person Standpoint

Stephen Darwall claims that we are beings morally constrained because of our capacity of address. He holds address is determined by the form of morality, and if he is right, then he must succeed in explaining the form of moral obligation - the fundamental principle of morality - deriving it from the form of address.

In *Moral Obligation: Form and Substance*, we find a good synthesis of Darwall's ideas. He puts his point in the following way:

I argue that moral obligation is conceptually tied to moral responsibility, specifically, to a fundamental answerability we have to one another (and ourselves) as representative persons or members of the moral community. Who are this 'we' and 'us'? Anyone, I argue with the psychic capacities necessary to enter into relations of mutual accountability. I call these capacities second-personal competence (...) Moral obligations are what we as second-personally competent persons are answerable for to one another (and ourselves) as representative second personally competent persons. This, I shall say, is moral obligation's form. (Darwall, 2013, p. 40)

3.1. A reading of Kant's Categorical Imperative

The ideas of moral responsibility and moral obligation as determinants of the form of address come from Darwall discussion of Kant's arguments in the *Groundwork* and in the *Critique of Practical Reason*³⁰ in his former book *The Second Person Standpoint*.

In the *Groundwork* Kant proves that morality's fundamental principle is the Categorical Imperative (CI) and it holds **if, and only if**, the will has the property of autonomy. Kant's aim is to prove the point mentioned in the Preface to the Second Edition of the First Critique (B XVIII), where he claims that moral obligation is conceptually tied to moral responsibility, for we can only

³⁰ Darwall claims the recognition of the other's authority to be Kant's strategy in the *Second Critique*. We can't rely on the argument that attempts to prove that autonomy is an inescapable assumption of any possible practical reasoning. Belief formation is not independent of facts. We don't decide whether or not to believe the facts we are presented to. We respond to independent objects and states of the world when we form beliefs.

be responsible if we are in a sense autonomous free agents. The problem after Kant's analysis of moral conceptual connections in sections 1 and 2 is to show that after discovering the principle of morality, it is necessary to show its empirical validity. However, in order to show the application to beings like us, "[...] it is insufficient to establish this biconditional (that the CI holds if, and only if we are autonomous), since it is merely an analytic truth that is consistent with morality's and autonomy's both being mere "figments [of the mind]" (Darwall, p.30).

Accordingly, in *Groundwork* 3 Kant claims that autonomy is a necessary presupposition of the practical standpoint -- as he says "we cannot act otherwise than under the idea of freedom". Hence, we presuppose we are autonomous beings in order to reason and judge how to act and this commits us to the Categorical Imperative. Although it looks like a solution, it has problems. The mere presupposition of freedom is insufficient because we do not necessarily reason in practical terms making use of it. Darwall explains what is the kantian inspiration and why it fails.

According to Darwall, Kant argues that any judgment, whether practical or theoretical, must assume a freedom of reason at its base. Even though Darwall takes this to be undoubtedly true, "it does not follow that the reasons on the basis of which we judge are not themselves drawn from features of our judgment's object (and hence, in practical deliberation, from properties of the objects of our desire and volition)". Objects from reality compel judgment when we reason about what to believe. However, we are clearly not automatically forced to believe our experience of reality, otherwise there wouldn't be mistakes. We believe our experience of reality instead of desire or dream, because, as free agents we discern our best beliefs to be about reality. Theoretical reasoning involves a form of freedom that is analogous to autonomy of the will, Kant says, for as much as reality compels belief, obligation compels action. Nonetheless, Darwall argues that autonomy need not be assumed in any intelligible practical reasoning. He asks us to

[c]onsider what practical reasoning would be like if it were structurally analogous to rational belief formation about the empirical world. [...] A naive theoretical reasoner takes her experience as evidence of states of an independent world. An analogously "naïve" practical reasoner might take her desires and other forms of practical experience like pleasure and pain as forms of epistemic access also, only this time, not to the world as it is, but to how it would be good or bad for the world to be, how it should or should not be. (Darwall, 2006, p.31)

Darwall asks us to imagine a naïve agent with a desire that p based on past experiences of pleasure and pain. For him, it will seem that the world should be such that p is the case. So, if

p is good to our agent's eyes, then there is reason to bring p about. In a naïve practical reasoning no autonomy of the will has to be presupposed for the agent to desire a state of affairs and make it happen. He can see reasons to act purely based on features of the objects and events he has previous experiences. Hence, Darwall's point is that Kant's *Groundwork* strategy fails. Any argument that tries to state autonomy as an inescapable assumption of any possible practical reasoning assume morality's supremacy and necessity, but if no autonomy is necessarily presupposed in some practical reasoning, then morality cannot be taken as a consequence of the necessary way we think ourselves as agents.

The second Kantian strategy is presented in the *Critique of Practical Reason*. There Darwall summarizes Kant's argument as "the abandon of [Graundwork's] strategy and argues that the awareness of autonomy comes itself only through the recognition that we are bound by the moral law (the fact of reason)" (Idem. p.213). If the recognition of being bound by the moral law is a purely first personal awareness, then it is necessary to explain how this achievement is possible. In Darwall's opinion this is not possible by a first personal standpoint alone, there's an irreducible second-personal element in morality that cannot be derived from first person alone. To state briefly: "Moral obligations do not simply purport to provide supremely authoritative reasons. They are also what we are responsible for doing, what members of the moral community have the authority to demand that we do." (p.214)

Darwall will claim moral normativity to be essentially second-personal, present in any second-personal address or acknowledgment that commits us to another equal, to another person. Acknowledgement commits both, the other and me, to be sensible to emotions, to the practice of mutual justification and to act according to norms and reasons shared due to our second personal authority.

When we acknowledge the summons of another free and rational agent, we confront, in effect, the "fact of reason". [...] We presuppose the equal dignity of rational beings and our ability to act on a "law" or reason—a second-personal reason grounded in this dignity—that derives not from the value of any state of affairs or outcome that might be the object of a desire, but, ultimately, from what it is to be one free and rational person interacting with others.

What forces an assumption of autonomy of the will, therefore, is the second personal aspect of moral obligation, that is, that what is morally obligatory is what we are responsible to one another for doing. (idem. p. 33)

The consequence of this shift of strategy to state moral normativity is the need of an accurate analysis of address, since second-personal competence is needed in order to act morally. Now I will follow Darwall's argument that connects these two concepts.

3.2. Moral Obligation and Address

Now I would like to connect moral obligation and address. After presenting Darwall's reading of Kant's arguments, we can conclude that responsibility does not come from any a priori notion as "good will" or from the concept of "autonomous self-conscious being", but it comes from the concreteness of our psychic capacity of addressing others. How does that changes the image we have about moral concepts?

A first consequence of this view is that responsibility is not to be lived in an individual dimension only, as if responsibility were only a form of accountability to one's own consciousness. This concept, here, also relates to the dimension of mutual accountability, where I make myself responsible for others in dialogue, given our attunement and agreement -- as Darwall says we have a fundamental answerability to one another because we are potentially representative of a community. So, I have to make myself answerable to others' reasons in such a way that, If the reasons I believe to justify my actions are weak, then I can acknowledge other's reasons as better and stronger than mine. Hence I can rationally accommodate and conform to others' reasons so as to make them my own. Moreover, adoption of other's reasons is not only an act of intelligence and mutual understanding, it is also an aspect of responsibility for what we have been calling our "human dignity", given that preserving in one's acts the rule of reason is equivalent to holding oneself responsible for human values. From this it's not hard to see why responsibility entails duties to myself and duties to others, for whoever acts in the sense of responsibility, acts in ways so as to preserve what is common to me and others.

Once we acknowledge that our moral practices are pervaded by our capacity to be answerable to others, it seems natural to ask "*how address is morally determined?*", since it is not an obvious fact that address is a way to understand the sources of moral obligation.

First let's take a look on how second personal competence manifests. Others can ask for our attention, make promises, blame us, ask for holding ourselves responsible for our own actions, demand things by talking to us, by crying or simply looking at us in ways we know are full of resentment. All the possible forms of address are demands that presuppose our capacity to answer to an interlocutor, and answerability to these demands is possible because we can constitute two forms of relatedness: *de facto* and *de jure* relations to others.

Accordingly, address is fully expressed by humans when we act in ways so as to meet *de facto* and *de jure* conditions, for not all acts of speech or gestures result in address. To give

you an example, if you declare a man and a hippopotamus married, they won't be married. For, even if you are in the presence of an all white-dressed bride hyppopotamus, if you pay for a huge party and invite yours and the hyppo's family, you can't truly marry an animal who's not able to give you consent. Marriage needs consent, and it is a condition for marriages. Hence, you won't meet what Austin called felicity conditions³¹.

Felicity conditions may vary according to which kinds of address one is to perform. For example, in an act of military command there must be a commander and someone to obey the command -- this is not only a de facto condition, but also a de jure one. Each one must know of one's own position in relation to the other, behave accordingly and so on. So, a command, like any addressed speech act, must meet felicity conditions in order to obtain efficacy. Therefore, felicity conditions are the norms by which we recognize successful forms of address. What all this has to do with ethics? I quote Darwall

When someone attempts to give another a second-personal reason, she purports to stand in a relevant authority relation to her addressee. I shall say that her address *presupposes* this authority. By this, I just mean that her having the authority is a necessary condition of the validity of the reason she purports to address and is thus a normative felicity condition of successfully giving her addressee the reason. *Qua* attempting to give her addressee the reason, therefore, she must assume this authority, as she must assume the satisfaction of any normative felicity conditions of giving the reason.

In addition to the specific presuppositions carried by different specific forms of address, a major claim of this book is that second-personal address has certain presuppositions built into it in general. To enter intelligibly into the second-person stance and make claims on and demands of one another at all, I argue, you and I must presuppose that we share a common second-personal authority, competence, and responsibility simply as free and rational agents. (Idem. p.4)

But how the presupposition of authority really works? Some reasons are subjective and others won't feel compelled to accept them, other reasons are based on a sort of authority that is felt as too distanced or too general; they don't seem to mobilize anyone to act on the basis of conviction. If authority is to be recognized, then how does that work?

³¹ Austin, *How to do things with words* (1975)

3.3. Agent-relative and Agent-neutral

When two people make use of second-personal competence they have to make use of their authority over claims and actions that may be the content of the act of address. The content of address must fit our practice of mutual accountability being sufficiently objective, but also sufficiently subjective. This difference is expressed in reason giving practice as agent-relative or agent-neutral.

To better understand the difference, Darwall's example is this: suppose you are trying to give someone reasons to stop causing you pain, to remove his foot from on top of yours. In an agent-neutral perspective, you would try to cause the other's sympathy for your pain giving him reasons that do not concern specifically *your pain*, but anyone's pain. For him to want to relieve your pain, he would have to understand your pain as something objectively bad and not bad because it is caused specifically in your body; something he does not want for you because anyone's feeling pain is a state of affairs he does not want, for as any other rational agent he wants anyone to be free from something he understands as bad. In this case, he as much as any other agent would be in position to understand your reasons. On this agent-neutral perspective two features are very important:

First, in pointing to the reason, you would be directing him epistemically rather than practically, albeit on a question of practical reason. *Qua* this form of reason-giving, you would be asking him to agree, as it were, that there is a reason for him to do something rather than asking him to agree to do it. Any claims you might make would thus be on his beliefs about practical reasons and not directly on his will. Second, your being able to give him the reason would not depend in any way on his seeing you as trying to give it to him or as having any competence or authority to do so. Anything you might do to get him to see the reason would serve. It might be most effective, indeed, if he were to see you as so defenseless and vulnerable as to be unable even to reason with him, like a young child. (Darwall, 2006 p.7)

The other possible perspective is the agent-relative perspective. In this case, you assert your authority to demand that he stops causing you pain. Darwall says that when you make use of your authority, whether as a member of a moral community -- you want him to stop the action because you have the right to demand -- or because it's your foot and you don't want it to be hurt, then your claim is agent-relative. As Darwall says

Whichever, the reason you would address would be agent-relative rather than agent-neutral. It would concern, most fundamentally, his relations to others (and himself) viewed from his perspective within those relations, in this case, that his keeping his foot on yours causes another person pain, causes inconvenience, and so on. The reason would not be addressed to him as someone who is simply in a position to alter the regrettable state of someone's pain or of someone's causing another pain. If he could stop, say, two others from causing gratuitous pain by the shocking spectacle of keeping his foot firmly planted on yours, this second, claim-based (hence second-personal) reason would not recommend that he do so. (idem, p.8)

The difference between the agent-neutral and the agent-relative reasons to stop causing pain can be also respectively understood as reasons that have independent grounds of acceptance and reasons that have personal grounds. On the one hand, independent grounds move agents to act according to non-directed duties -- duties whose source is reason itself, but are not owned to anyone in specific. On the other hand, personal grounds move agents to act according to directed duties -- duties that sources are other persons taken as self-determining beings, hence duties that are owned to others. In the act of address it is expected that I give second personal reasons based on my authority, the authority that here comes from the fact that if it wasn't for our second-personal relatedness the demand would not exist. That's why we understand persons in directed duties as self-originating sources of claims³², for they can legitimately restrict other people's freedom, while in indirect duty the reasons to act are independent and no address is necessary to motivate the person to act. So, given this distinction, Darwall concludes that

A second-personal reason is one whose validity depends on presupposed authority and accountability relations between persons and, therefore, on the possibility of the reason's being addressed person-to-person. Reasons addressed or presupposed in orders, requests, claims, reproaches, complaints, demands, promises, contracts, givings of consent, commands, and so on are all second-personal in this sense. They simply wouldn't exist but for their role in second-personal address. And their second-personal character explains their agent-relativity. As

³² As John Rawls puts in *Kantian Constructivism in Moral Theory*

second-personal reasons always derive from agents' relations to one another, they are invariably fundamentally agent-relative. (idem. p.8)

A difficulty here arises when we try to connect agent-relativity and moral community. If agent-relativity is based on the authority a person has to claim reasons, it seems that too personal and idiosyncratic requests, claims, complaints, or demands can be insufficient or fail to justify to the eyes of a community. The addresser can give the wrong kind of reason to her addressee. Suppose someone asking for pain relief says "Stop hurting my foot! I don't want my foot looking ugly, for I want to wear sandals". Even though the kind of reason given constitutes a logical connection with the first claim, it is of the wrong kind if my aim is to convince someone that feeling pain is bad. The reasons we expect someone to give when asking for pain relief are of the moral kind, because what one wants is to compel morally the other subject. As we claim the other to reason morally we ask the other to recognize herself as a member of the moral community -- which is from where I base my authority after all. But what is the role of the moral community in this specific one-to-one relation?

3.4. Second Person and Moral Community

Roughly, Darwall says that my authority embodied in addressed speech or action is agent-relative, which is to say that it is derived from the agent's relations to one another. In addition to this, he has said in the beginning of this section that one's authority is present when one speaks as representative member of a moral community. So, if in order to share a normative point of view people must make use of second person competence, then it is as a member of a moral community that moral demands acquire force. Hence, moral obligation and moral responsibility are embodied in address when people recognize each other as member of the moral community -- as he puts it, when people adopt a first-person plural perspective of the moral community.³³ Moral community works as the common grounds to which both persons can mutually refer.

Now suppose you're hiding Jews in your basement during the world war II. One day a Nazi soldier comes to your house saying that he has seen suspicious activity in the neighborhood -- people buying more food than needed, water consumption has reached high levels strangely, etc. So, he asks whether you're hiding people in your house. If you're to answer the Nazi soldier in terms of the actual moral community, the Nazi moral agreements are certainly

³³ Darwall does not say explicitly whether his conception of second-person standpoint accords to Rödl or Peacocke. All he says about pronoun analysis is : "It is the perspective one assumes in addressing practical thought or speech to, or acknowledging address from, another (whether as an "I" or as part of a "we"). [...] It involves practically directed and directive thought, thought that is addressed to, and that makes a claim on, a free and rational agent." In my opinion, this places him closer to Rödl joint action conception of acknowledgement.

not the ones you want to take as yours, since you're the one who is trying to protect innocent people. So, how the appeal to moral community works? How moral community is related to address? How the promised moral obligation built into address is present in thought? Darwall's solution to these difficulties is to think the act of address as directed to a virtual community:

[...] we might think of moral demands as being 'in force' if members of 'the moral community' are prone to make them. But the moral community as I understand it is not any actual community composed of actual human beings. It is like Kant's idea of a 'realm of ends', a regulative ideal that we employ to make sense of our ethical thought and practice. So, as I am seeing it, it takes neither an explicit actual demand nor a demand that is implicit in actual human beings being prone to make it, either individually or collectively, in order for a claim or demand to be in force. The demand is made by the 'moral community' and by all of us insofar as we are members. (Darwall, *The Second Person Standpoint*, 64-65)

Darwall thinks address in counterfactual terms -- not only as a one-to-one activity, but also as a one-to-many, for when you face another person in address with moral content, you are not only speaking to the person in front of you, but to the person as a representative of the moral conventions of an ideal community, a community where everyone is treated as an end-in-itself. Hence, even though telling the truth is a duty that I have as rational agent and in relation to another rational agent, when telling the truth takes others as means to keep myself alive, then I'm violating other people's dignity and autonomy.

As ends in themselves, all persons can speak as representatives of moral community, for they constitute the realm of ends -- the realm of beings with the capacity for autonomy or self-determination. This makes all subjects responsible for holding the meaning of their practices either as part of a community or as an individual -- this last aspect is what makes them all possible authors. Every responsible subject also has authority to try new valid steps in the moral agreement, as well as authority to recognize steps given by others.

3.5. Derivation of Moral Principle

What we have to answer now is how address is constituted by the moral principle (CI). So far, we have agreed that address puts both poles of the relation in mutual accountability, which entails that both persons are in equal position of authority and membership to the moral

community. Now we want to understand the derivation of the moral principle from the act of address.

Moral obligations we have to others entails that others have rights against us, for we are obliged to respect others. If we owe to others respect for their rights, this means things can go wrong, for I can fail to regard moral obligations and then I can awaken reactive attitudes like indignation or moral disapproval.³⁴ These reactions are the ordinary symptoms that we experience when norms are not followed correctly, so when someone wrongs me, I hold this person responsible and blameful. Darwall follows the clues that reactive attitudes leave behind to start the derivation of what he believes are the substantive principles of morality, the principles that give us the form of moral obligation in general.

The first step is to distinguish between what morality recommends and what morality requires as obligatory. As Darwall says, there is "a conceptual gap [...] between the ideas of, on the one hand, what morality recommends, however strongly, and, on the other, what it demands or requires." (p.43) This conceptual difference is seen when we think of the openness that a recommendation leaves for the agent. If morality recommends x doing ϕ , then the reasons for doing ϕ are categorical. Something recommended might not be done, no one is blameworthy for not acting in the recommended way. So, it's in the agent's hands to do or not to do ϕ .

Moral obligation is quite different. Any moral obligation that requires you to do ϕ must be taken as a moral demand. If a person fails to do what is morally required, then it seems that in some sense this person is rationally criticizable or blameworthy. In Darwall's words "Moral obligations are moral *demands*, and moral demands are whatever we are legitimately *held* to or answerable for, that is, where violations are *blameworthy*, the agent lacks a valid excuse" (Darwall, *The Second Person Standpoint*, p.43).

His second step in the derivation of moral principles attains the implied features of the blamer and the blamed. Darwall holds that it is "[...] the conceptual connection to accountability and blameworthiness that explains moral obligation's overridingness and not the other way around. If an action is morally obligatory, then it is blameworthy not to do it without excuse." (p. 44) The argument Darwall defends is similar to what Bernard Williams and many others have claimed. Moral blame and responsibility are in a par with epistemic responsibility. Analogously to Moore's argument that someone who asserts that p , believes that p ; for it makes no sense to assert something and deny that one believes in it; "so also it does not make any sense to blame someone for doing something and then that he had, nonetheless, sufficient reason to do it, all things considered" (p.28)

³⁴ Here Darwall explicitly follows Strawson's *Freedom and Resentment* arguments already presented in chapter 1 about reactive attitudes.

To blame someone is in some way to participate in interpersonal relationships. Blame and resentment are moral reactions that have addressive character even when not expressed. In the very act of address, I must assume that the other has the capacity to address him/herself, otherwise, if I don't take the other as capable of understanding and holding himself responsible, that would be a performatively self-contradictory act³⁵. To see the point Darwall is trying to make, think that someone has wronged you. So, your natural reaction is to blame this person for doing something wrong. But blame only makes sense if it holds the blamed person responsible for her/his action. Naturally, we expect the blamed, as a member of the moral community, to think of his justifications for doing so. If she agrees, she will think of herself as blameworthy. Otherwise, she will try to account her action to me and she will try to say that she was obliged to do so for whatever reasons. Obviously, if she was obliged, then she is not blameworthy. In both cases what is essential is that she was taken as a free agent, that is, she is taken as someone able to guide herself by self addressed demands -- she is able to take up a second-person standpoint on herself, as Darwall says.

Moral reactions can be classified in two groups according to Strawson. First, we have the personal reactive attitudes, like resentment and guilt, that are felt as if from the perspective of a participant in the events giving rise to the moral attitude. And then we have impersonal reactive attitudes such as indignation and moral blame, that are "felt as if from the perspective of a third party's point of view."

The reactive attitudes I have so far discussed are essentially reactions to the quality of others' wills towards us, as manifested in their behaviour: to their good or ill will or indifference or lack of concern. Thus resentment, or what I have called resentment, is a reaction to injury or indifference. The reactive attitudes I have now to discuss might be described as the sympathetic or vicarious or impersonal or disinterested or generalized analogues of the reactive attitudes I have already discussed. They are reactions to the qualities of others' wills, not towards ourselves, but towards others. Because of this impersonal or vicarious character, we give them different names. (Strawson, 1962 p.15).

Both impersonal and personal reactive attitudes must be interpersonal, but the difference between them is the actuality of personal connection that enables expression in the

³⁵ Darwall claims the recognition of the other's authority to be Kant's strategy in the *Second Critique*. We can't rely on the argument that attempts to prove that autonomy is an inescapable assumption of any possible practical reasoning. Belief formation is not independent of facts. We don't decide whether or not to believe the facts we are presented to. We respond to independent objects and states of the world when we form beliefs.

first case, and its absence in the second. The personal reactive attitudes are only possible if there's a connection between people, for one cannot resent or forgive someone if there's no actual connection between them. On the flip side, impersonal attitudes involve the same second-personal element present in address, as in personal attitudes, but differently from it, this kind of attitude is taken from the perspective of a *representative person* rather than any individual's standpoint.

Analogously, two different sorts of authority will emerge from the two different sorts of reactive attitudes. *Representative authority* is the general kind of authority connected to impersonal reactive attitudes and *Individual Authority* is the specific kind of authority connected to someone who's hold answerable to resent wrong actions. Representative authority emerges when we disapprove or blame someone who violates specifically a moral obligation. Someone who answers demands according to the point of view of Representative Authority is someone who is not caught in any idiosyncrasy. We don't demand of anyone on the basis of representative authority as individuals with personal traces or specific history, we demand of them as representative persons, from the impartial perspective we and others share, as if free from individuality. In contrast to representative authority, individual authority emerges when someone has wronged *you as an embodied agent*. As Darwall says, when individual authority is called upon, then "you have a distinctive individual authority as the victim to hold him answerable, for example, to resent the wrong, which you can exercise or not at your discretion. You can seek compensation or not, forgive or not, and so on." (p.46) so as someone in the position of an individual authority, it's in your hands what to do when someone wrongs you.

As Darwall sums up:

Warranted impersonal reactive attitudes, and thus moral obligation period, implicitly involve representative authority; and warranted personal reactive attitudes, and thus obligations to and correlative claim rights against, implicitly involve individual authority. (Darwall, 2006.p.46)

Now we are apt to reach Darwall's third step: the derivation of the substantive principles of moral obligation from second-personal competence. Darwall takes second personal competence to involve not only reactive attitudes, but in addition to it he adds imaginative and empathetic abilities necessary to think counterfactually. Putting yourself in someone else's shoes is in his understanding the best way to access others as representative persons, for seeing the world from the other's position enables us to see a better balance between specific and general. Once you reach the perspective of a representative person, you understand which demands would be warranted from this perspective. This capacity to switch between one's own and other's positions with further addition of the capacity to regulate oneself by normative

judgments about conduct is both sufficient and necessary for moral obligation, for they constitute the second person competence.

But what could warrant that any person would, merely from second person competence, be in position to understand her moral obligations in the same way as any other person? On the universalizability requirement of morality, he says that

We have been assuming that what matters for moral obligation is what demands would be warranted from the impartial perspective we and others share, namely, that of representative person. Moreover, if others are bound by the very same fundamental moral obligations we are, then others must be able to see these demands as warranted from the very same perspective also. It follows that something can be a fundamental moral obligation, then, if, and only if, it involves a demand of anyone (i.e., any second-personally competent person) that anyone can see to be warranted from the perspective of a representative second-personally competent person. (idem. p.49)

As we have seen before, Darwall believes that from the second person competence, we can reach the concept of representative person. In his account, second person competence is what connects us to moral obligation, since moral obligation entails a general demand, this in its turn entails that it's only as representative persons that we meet demands that warrant the possibility of accessing an impartial perspective. This impartial perspective is the basis from which we can see moral obligations as duties to a person *in general*. Now he says that every representative person must be able to see *the same demands* as warranted by the same perspective. But which perspective is that? His answer: the perspective of free agents. He says "[o]ne cannot coherently hold someone responsible for (freely choosing) to do something while simultaneously undermining his ability to make the choice himself or otherwise usurping his agency." So, the second person competence is what leads us to the discovery of ourselves as free agents -- this is why other's will is the source of the Categorical Imperative in Kant's Second Critique.

Darwall is saying that in a practical sense, thinking other minds is the same as to think a rational free agent and the way I recognize another mind is the same way I go from my reactive attitudes towards a conception of responsibility of the other as a representative person. In that sense, others are not merely taken as valuable objects concerning which we make choices, but our grasp of others is in the mode of subjects to whom our choices are in some way addressed.

3.6. Objections

Darwall thesis that we are moral beings because we have second person competence conflicts with the idea of counterfactual deliberative thought when his theory reaches full development. First, Darwall claims that an adequate account of what is to apprehend the subjectivity of another person should explain the kind of moral (non-instrumental) significance that other person have for us. As Korsgaard puts in her answer to *The Second Person Standpoint, Autonomy and the Second Person Within (2007)* "Suppose someone does not hold herself or anyone responsible for meeting demands, or resent it when they do not?"³⁶ Even though there might be psychological pressures that lead us to take a second personal standpoint, there isn't any incoherence in avoiding the second-person standpoint altogether -- and even Darwall acknowledges that. The way he constructs his argument, makes morality dependent on psychological pressures we should feel that fail to account for the overriding character of moral obligation. It seems that moral obligation comes from experience of emotions, but if this is so, we cannot reach the universalizable and necessary character required by morality. Korsgaard says that there is nothing in morality that is necessarily connected to the capacity of having a second person standpoint -- for nothing coming from reactive attitudes can rule moral obligation. In addition to this, if we are to think morality counterfactually, why do we need the second person standpoint? Do we really need to achieve felicity conditions in second personal address to experience reactive attitudes and then build all the notion of counterfactual thought and empathy?

Another strange feature in Darwall's theory is pointed out by Korsgaard concerning the discovery of one's own autonomy. She says in the same answer to Darwall's book from 2007 that

On Darwall's account, the second-person standpoint does not seem to be unavoidable, the way the standpoint of first-personal deliberation is. If we are autonomous, then the person who declines to take up the second-person standpoint fails to know something intimate and important about his own agency. But unless he has a reason to take up the second-person standpoint and its presuppositions, it is possible that he will never know. And that conclusion can be generalized: if it were not for the fact that it is so psychologically difficult to avoid holding one another accountable, it seems as if we might never have discovered that we are autonomous. But then how could we ever have been autonomous? And how do we go about "presupposing" the autonomy of

³⁶ Autonomy and the Second Person Within: A Commentary on Stephen Darwall's *The Second-Person Standpoint* Christine M. Korsgaard.(2007)

others in second-personal address, if we have no independent access to this feature of human agency? (Korsgaard, 2007, 22-23)

Since to Darwall the discovery of autonomy is dependent on the second person standpoint and the latter is contingent, then judging myself as a free agent is also a contingency. This is very controversial, for if I am autonomous, then I must have some experience of my own autonomy. If I'm capable of acting according to reasons, then I at first necessarily know that I am autonomous. It is hard to accommodate the necessity involved in those concepts with Darwall's theory of address and morality.

Given that we have given Korsgaard some voice as Darwall's interlocutor, now we will examine her understanding of the relation between recognition and morality.

4. Korsgaard's Alternative

Korsgaard holds that the source of the normativity present in recognition is self-consciousness. We unavoidably have a second person standpoint over ourselves because we are in the first place answerable to ourselves. This fact, she says, "is made clear to me by the voice of the second person within" (Korsgaard, 2007, p.23), which is the voice one listens when interacting with oneself. Ultimately, her project is to support the view that "[...] some duties really are owed *to* others, [...], in much the same way that we may be obligated by ourselves." (Korsgaard 1996, 134)

In what follows, I will account for Korsgaard's understanding of our self-conscious nature, which in her view is the basis of morality. She claims action is self-constitution, for I constitute my unity as an agent through acting. Besides, she acknowledges that self-constitution involves self-address, because when I address myself I speak as the person who is the ground of the unity of my actuality as an individual agent. Self-address in its turn is an act of obligation to oneself, for it involves duties one has to oneself. Finally, in connection to this I will examine her claim that what holds in my self-relations also holds for my relations to others.

4.1. Self-Consciousness

Korsgaard starts with the fact that we can step back from what we are tempted to think and do. She claims self-consciousness is what makes possible the reflective relation to ourselves, for self-consciousness is a reflective structure that places us in relation of authority over ourselves. Authority is based on the ability to access the potential grounds of beliefs and actions and in the capacity to act in accordance with these very grounds. Authority allows one to account one's own acts and reasons to oneself. But if, as an author of myself, I am also responsible for sticking with my beliefs and grounds for actions, then authorship comes with

responsibility. So, in a sense responsibility leads me to step back from what I am tempted to do, allowing me to ask myself whether I should do that and whether I have good reasons to do so.

Once I step back from what I was tempted to do and affirm other possible actions in its place on the basis of reasons, I act rationally. In order to do that I need a principle that enables me to think my actions as rooted in reasons: I need a law or maxim. As Korsgaard suggests

When I am conscious that I am tempted to do something because of something else, I can ask myself whether I should do that, and this amounts to asking whether the consideration on which I propose to act provide reasons. To answer in the affirmative is to adopt a certain maxim of action as a law governing my conduct; [...] Thus I act under my own authority as a lawgiver, and I am accountable to myself if I do not. So my reasons [...] are grounded in the authority the human mind necessarily has over itself. [...] I think that every rational agent stands in what Darwall would call a second-personal relation to herself – she has a second personal voice within” (Korsgaard 2007, 11).

If Korsgaard is right, then my responsibility does not come from any primordial bipolar obligations I have to others. By itself, the structure of self-consciousness does not get us to the idea that I owe a certain kind of treatment to others. Instead "it only gets us to the idea that I owe it to myself to treat you in a certain way." (Korsgaard 2007, 11). I know I owe to myself to treat others in certain ways because to act morally is to act in accordance with the unity that self-consciousness demands in its self-constitution. Self-consciousness permits one to know morality from the capacity to interact with oneself, from the capacity of weighing reasons and acting in accordance to them. Nonetheless, this capacity is constituted and its constitution is the constitution of the unity of the self. I hope we can make things a bit clear in what follows.

4.2. Self-Constitution Argument

Here I will side with Korsgaard for what she calls the argument for self-constitution. This is supposed to show that the way we interact with ourselves is the way we can interact with others.

First, Korsgaard states that humans are condemned to choice and action. We cannot fail to act, even when we decide to stay immovable, for this is still in some sense an action. Action and choice are pervasive, they are presented to us as necessary since we are self-conscious. Hence, the nature of this condemnation is the necessity we have to face, the necessity that reason impinge on us.

Human action³⁷ is the articulation of perception, intentionality, attribution and purpose. Two of these features are fundamental to understand the specificity of the human agency: Purpose and Intentionality. Purpose entails efficacy, which is guided by the Hypothetical Imperative -- the principle that permits us to think the means to achieve the ends of our actions. Intentionality entails "a movement of an animal that is guided by a representation or conception that the animal forms of his environment" (p.97) . In this sense, wild animals are capable of intentional states, for they are agents capable of inserting themselves into the causal order. However, their conception of their environment is loaded with the significance that biological instinct provides. They do not choose action in the sense that they are interested in constituting their animality in a variety of ways. Their actions are merely voluntary. In this specific sense animals and humans are agents. However, human agency is reinterpreted by self-consciousness. This makes us autonomous and changes the nature of our agency, Human will is capable of self-determination, whilst animal agency is guided by the biological function.

Self-consciousness creates the need for the principles of reason, which are then more firmly separated from the associated incentives than their instinctual predecessors were. But self-consciousness also transforms the other side of the equation—it transforms incentives into what Kant calls *inclinations*. (Korsgaard, 2009. p.120)

As beings capable of action, we are agents. Self-conscious agency leaves its effects on us, for it carves practical identities out of our actions. Agency constitutes in our personality a long stratification process of personal identities that we assume, some as naturally belonging to us, some as crafted and chosen by us. These practical identities combine in each of us in different ways, they give us different sources of motivations, and with them different sources of obligations. Some of us will be parents, others will dedicate their lives to science, others to nature, others will mix things up and so on. As we value ourselves as bearers of contingent practical identities that give us reasons to act, we constitute our different personalities. Having a personality here means being a person. A person is someone who constitutes oneself in a unity of reasons and actions. So, being a person is to be someone in particular, and in order to do so, one must commit oneself rationally to the motivations that the practical identities demand of us.

Here the understanding of what is to be a person with personal identities and the concept of recognition blend in. If we are to understand the possibility of having relations of recognition between persons, we must understand how sameness and difference compose this

³⁷ In phototropism, a plant turning towards the sun is not an action. Even though it is a reaction to the environment, it is only governed by perception, but not by any representation of the environment.

relation. For we are all different and unique in some sense, and we are all persons and equals in a different sense. Darwall's bet is that we are capable of assuming the positions of representative person and from this position speak as a maximally rational agent. Korsgaard believes our form of life have the specificity of openness to differentiation and individuality, so we don't need to situate ourselves in any abstract position to reach the second person standpoint. We must constitute ourselves throughout life, where living is the activity of self-constitution. But in order to constitute oneself, one must act according to constitutive standards that rule different kinds of personal identities.

Personal identities are ends to our actions. We constitute them in order to constitute ourselves. Of course, not all action virtuously reach its end, for efficacy comes with the development of virtue. However, as much as a badly constituted house still is a house, a badly constituted human still is a human. We inherit from education and culture rules and norms that help us in constituting the standards of our personal identities. So, a if you are willing to be an athlete, you must train, eat healthy and sleep well. The normative standards rule your actions towards the proposed end, not as causes, but guiding³⁸ you to conform to your ends. The unity of these actions teleologically constitute the personal identity of an athlete. So, if to act is to determine yourself to be the cause of a certain end, as even animals do, then to act self-consciously is to conceive of yourself teleologically as the first cause of a certain end.

One must be able to attach the "I do" to the action in the same way that, according to Kant, one must be able to attach the "I think" to a thought. [...] And I also believe it is essential to the concept of agency that an agent be unified. That is to say: to regard some movement of my mind or my body as *my action*, I must see it as an expression of my self as a whole, rather than as a product of some force that is at work *on* me or *in* me. (Idem. p.18)

Korsgaard argues that to unify action and motivations in a rational way, there must be a principle that relates particularity and generality. For the specific blend of personal identities that each person carries with herself and the rational structure that underlies each agent's motivations need a unifying principle which is also a formal principle. Korsgaard suggests the Categorical Imperative is the principle that unites the specificity of each person with the generality that norms entail.

The Categorical Imperative, as we have seen before, is the formal principle that permits the formulation of a maxim with universalizable content, but also permits agents to treat

³⁸ Korsgaard believes the norms and ethical principles guide action not as premises that we apply in deliberation, in the same way that modus ponens are not premises of inference. Ethical principles are principles of rational action, principles we embody in action. Hence, knowledge of principles is in a sense self-knowledge (Self-Constitution, p. 67)

themselves as ends with specific lives. So, whoever wills an end is a self-constituting agent, for one is the cause of that end. Nonetheless, whoever wills the ends, must will the means -- which in this case is oneself. So, if the Categorical Imperative is a principle that permits self-conscious agents to constitute themselves, then the means for that must be included in its formulation. Given that the Hypothetical Imperative is the principle by means of which we think means for ends, in the formulation of the Categorical Imperative the Hypothetical Imperative corresponds to the part where I pick myself, what motivates me, what awakens my feelings and emotions, as a cause of change in the world to constitute myself as an end.

We haven't yet accomplished the task of explaining why the Categorical Imperative permits one to blend particular and general. The Categorical Imperative is the law of acting only on maxims that anyone can will as a universal law. However, when it comes to determine an action, the motivations must be felt and thought in particular, as something that I, in a first person perspective, should endorse as rational, and as having to do with my life. So, the Hypothetical Imperative will take my life as I have lived it, my love for books, for Kant's ethical theory, or whatever. Then, the Categorical Imperative articulates one's personal identities and sources of motivations given by these identities, and the rational and justificational structure that informs and unites those identities in one self-conscious being. Korsgaard example is her desire to be writer of a book that is so good that it will be required reading in Universities. She shows us the wrong and the correct articulation between personal reasons and general interest:

Wrong structure:

- 1) I want my book to be required reading (private/individual reason)
- 2) Therefore: I shall write a good book

Correct structure

- 1*) Someone should write a book that is good enough to be required reading
- 2*) So, I want to be that someone. (idem. P. 210)

In the correct model of deliberation, (1*) is the public reason that can motivate anyone who shares with Korsgaard the taste for philosophy, ethics and Kant. In addition to this, (2) comes as the self-determination, where the first person pronoun will give the agent a place instantiating the public motivational structure. The conclusion is that our personal identities are constituted in ways that they can be willed as possible human lives, for they are not sources of motivations for only one person. They are not what Nagel dubs as agent-relative sets of motivations, but since these motivations are rationally structured, they are possible actions and identities for all humans that may want to inherit the path of other previous agents.

4.3. Interaction Argument

Now we are close to understand how recognition is possible in Korsgaard's account. She holds that interaction with another person / agent poses an analogous problem of unity to the one we observe for the individual action. In individual interaction, when one speak out loud, writes or goes to psychoanalysis, one must recognize one's own desires and motivations, so as to affirm or refuse them. In the same way, if interaction between you and me is to be possible, we must be capable of mutual recognition. The way this is possible is by means of what we can share -- we share reasons, we make them public. Here, Korsgaard takes Kripke's argument about the publicness of meaning as an argument that extends the feature of publicness to *all* meaningful activity. She says:

On the public conception of reasons I must take your reasons as my own. ... if personal interaction is possible, we must reason together, and this means that I must treat your reasons ... as *reasons*, that is, as considerations that have normative force for *me* as well as for you, and therefore as public reasons. And to the extent that I do that, I must also treat you as what Kant called an end in yourself – that is, as a source of reason, as someone whose will is legislative for me. (idem. p.192)

So, in personal interaction one constitutes oneself. One thinks one's actions (adopts maxims) by means of the Categorical Imperative and treats one's reasons as having normative force to all persons. If my reasons are reasons for at least one person, they can be thought by another person, their meanings are not private. Hence, they can fit other lives, they can serve as source of reasons for others. As a consequence, Korsgaard's account of agent-relativity and agent-neutrality differs from Darwall. She takes all possible agent-relative reasons to be a possible pathway for human life, hence necessarily universalizable.

[...] when you come to see that your contingent practical identities are normative for you only insofar as they are endorsable from the point of view of your human identity, you also come to have a new attitude towards your contingent practical identities. You come to see them as various realizations of *human possibility* and *human value*, and to see your own life that way: as one possible embodiment of the human. (idem. p.212)

Even though up to here we have a theory that enables us to understand how personal histories constitute possible ways of living one's humanity, this is a shift from singular first person to a plural first person. The relation from first person to second person hasn't been explored. But is it a necessary bridge? Can I avoid entering the second person standpoint with others and still be an agent? Can I avoid playing the language game of accounting myself to others and to myself and still constitute agency?

Korsgaard believes that interaction with oneself is necessary to become an agent. If action is necessary, I *must* interact with the "conscious inhabitants of my body" -- the so called identities I necessarily constitute --, that is to say: I must do something about my most basic needs, the natural demands that my self-conscious nature inescapably becomes aware of. So, if one cannot avoid constituting one's own identity, then one cannot avoid interacting with oneself.³⁹ I take this argument to establish further that my interaction with myself awakens me to the meanings displayed by body of others. When I act intentionally, I know from my own experience that such and such actions can constitute such and such identities. In interaction with myself, I learn to read others. It is as if others were texts. The body of others talk to me in address, they mean something even when they do not address me specifically or when they do not make use of words. Perhaps we cannot avoid interacting with others because I can be an other to myself -- in the sense that otherness is not only an experience we have coming from other people, but also coming from myself, from my own past or my conflicting identities. Hence, for Korsgaard the second person standpoint is not a source of the discovery of my moral obligation because I necessarily interact with others, but simply because I can't avoid interaction with myself.

Wittgenstein tells us that inner processes stand in need of outward criteria, but he does not say that outward criteria are apprehended only from a second person standpoint. We observe others at distance not only from second person standpoint, we can observe others third personally also, for we learn from other humans in distant observation. If Korsgaard is to agree with Wittgenstein's understanding of the inner life when she talks about interaction with oneself, she must acknowledge that such interaction is only possible because outward criteria is assumed as a means to constitute the inner dialogue. So, we cannot avoid some form of interaction with others, either interacting second personally or third personally.

³⁹ Autism is a psychological phenomenon discussed in many different areas, from psychiatry to psychoanalysis. There's no consensus about what it actually is, but one of the agreed consequences of autism is the incapacity for self-constitution as a person. So, when I say that one necessarily constitutes identity due to one's self-conscious nature, it is important to have in mind that some empirical conditions must be met in order to have a regular development.

Korsgaard's next step tells us that the basis of interaction with oneself is respect. She comments the example of Parafit⁴⁰

The story goes like this. The nineteenth-century Russian is now, in his youth, a socialist, and he plans to distribute large portions of his inheritance, later, when he comes into it, to the peasants. But he also anticipates that his attitudes will become more conservative as he grows older, and that he may not think that this is the right thing to do when the inheritance is finally his own. So he makes a contract *now*, to distribute the land when he gets it, which can only be revoked with the consent of his wife, and he asks his wife to promise not to revoke it *then*, even if he tells her *then* that he has changed his mind, and that she is released from the promise. (Idem, p.160)

So, the question here is: Is it "ok" to simply change one's own mind about past promises? What is the point of making promises if there's always the possibility of changing who one is? On the one side it seems normal to tell others "I changed". On the other side it seems morally wrong to say "I changed, I don't care anymore". The Russian nobleman simply believes he is going to change, but he asks his wife to prevent him from not doing what he deserved when young. She also must promise him she won't let him break his promise, but "then, Parfit says, "It might seem to her as if she has obligations to two different people." (p.160) So, respect here enters as a way to relate to your own history, to your past identities and to the unity and integrity one constitutes. My integrity and rationality are not two separate aspects of my personhood. Korsgaard asks us how should we treat the nobleman who breaks the promise he made to herself, like this noble socialist. Her answer is: you should not treat a disunified person as a unified person, or in other words, don't treat her as a person at all.

Unity is constitutive of personhood, for in interaction with oneself the principles of rational agency are supposed to unify the agent's will and parts of the soul accordingly to the human rational dignity. So, basically it is the lack of self-respect that is missing in the case of socialist's interaction with himself.

"I have suggested that the conditions for successful personal interaction are joint conditions of respect for the other's humanity, and the treatment of her reasons as considerations with public normative standing ... But action is simply interaction with the self. If this is so then respect for humanity in one's own person, and the consequent treatment of one's

⁴⁰ Parafit, D. Reasons and Persons, p.327-8

own reasons as considerations with public normative standing, are the conditions that make unified agency possible.” (Idem. 204)

So, what holds for my self-relation holds for my relation to others. As Korsgaard says, “every person interacts with others as he interacts with himself [...] Inward and outward justice go together” (p.206) to what she adds “Thinking [...] is just talking to yourself. And since meanings are public, talking is just thinking in the company of others” (Idem, 197).

4.4. Some Remarks and Objections.

One of the central arguments of Korsgaard is the interaction with oneself argument. We intuitively understand this idea by means of the example of the noble young socialist. However, if this is the second person standpoint one can take on oneself, how can I be unfair in interacting with myself? Does it make sense to say that I failed to respect myself? Korsgaard's picture of human moral life in some sense resembles Wittgenstein's image of the two hands giving a gift to each other.

“Why can't my right hand give my left hand money? – My right hand can put it into my left hand. My right hand can write a deed of gift and my left hand a receipt. – But the further practical consequences would not be those of a gift.” (PI, 268)

A first difficulty with Korsgaard's ideas show up when we try to understand what is for one to owe duties to oneself. This seems to fail to conceive one's moral life as requiring some necessary distance to oneself. If I wrong myself, it seems that I am already in position to punish me, but if I am in such a position I am also immediately relieved from my mistakes, since the one who's having the right infringed and the one who's restituting justice are one and the same person. So, if I wrong myself, I am also the only person immediately compensated by my own punishment. We can take a better look at this following Aristotle, when he claims in the *Nicomachean Ethics* (NE, 1138a19-26), “If an unjust action is voluntary, the doer of the action cannot be the same as the sufferer, since one cannot break into one's own house and steal his own property”. So, if Aristotle is right, there's certainly the need to explain how can one be unfair to oneself.

It seems that to solve this difficulty we cannot think the obligor and the obligated under the same sense, even though we should think them as having the same reference. In a position the subject is an agent, in the other the subject is passive. In an attempt to understand this I think the agent first under one aspect: a sensible being, who has a specific history, gender,

family, accidents, life-events, and so on and, while in another complementary aspect, the agent is an intelligible being, who believes, thinks, deliberates about the former aspect of his existence. This is a well-known Kantian stance, articulated by Kant himself as follows:

When a human being is conscious of a duty to himself, he views himself, as the subject of duty, under two attributes: first as a sensible being, that is, as a human being (a member of one of the animal species), and secondly as an intelligible being (not merely as a being that has reason, since reason as a theoretical faculty could well be an attribute P of a living corporeal being). The senses cannot attain this latter aspect of a human being; it can be cognized only in morally practical relations, where the incomprehensible property of freedom is revealed by the influence of reason on the inner lawgiving will.

Now the human being as a natural being that has reason (homo phaenomenon) can be determined by his reason, as a cause, to actions in the sensible world, and so far the concept of obligation does not come into consideration. But the same human being thought in terms of his personality, that is, as a being endowed with inner freedom (homo noumenon), is regarded as a being that can be put under obligation and, indeed, under obligation to himself (to the humanity in his own person). So the human being (taken in these two different senses) can acknowledge a duty to himself without falling into contradiction (because the concept of a human being is not thought in one and the same sense). (Kant, *The Metaphysics of Morals*, p. 215)⁴¹

In interacting with myself I am subject (author) and an object. As author I must try to unify my incentives -- the many identities I am/have (animal, philosopher, south american, male, etc) -- If I don't succeed in doing so, I threaten my integrity. Obviously it is necessary to say that my identities include many things I dislike, things I carry with me because they are rooted in who I am. Hence, I can wrong myself when I fail to respect my integrity, when I fail to see my own history as the product of both what I inescapably inherit from the past and how it relates to my own choices.

A second objection mirrors the first objection. The structure of self-consciousness makes some forms of interaction with oneself absolutely different from interactions with others. Who cleverly observes this is Matthias Haase in *For Oneself and Toward Another: The puzzle about*

⁴¹ Kant, *Metaphysics of Morals* (1797)

Recognition. He shows that the set of the possible reactions one can obtain when addressing others and to the possible reactions one can have when addressing oneself is surprisingly unequivalent. As Haase says

Of course, a person may be said to ask herself a question and proceed to inform herself about the matter. But no matter how much she addresses herself to it, there will be no second person within whom she tells or from who she learns. The grammatical subject of the sentence does not denote someone who already knows before the act, and its grammatical object does not denote a recipient of knowledge passed on by the former. The activity so described is therefore not a transaction of imparting knowledge. In its ordinary reflexive use, "informing" is not transitive. (reference)

So, here we must distinguish between "being capable of having a second person standpoint on oneself" and "having a second person standpoint on others". The first is closed to learning new information. The second is open to new information. If I address myself asking about the temperature on Mars and I know I never knew about such thing, then it's non-sense to ask myself that. Since the verb is used in the reflexive mode, we don't constitute a relation between agent and patient. With other forms of self-address it seems possible to ask myself, e.g., when I don't know what to do with my old papers from faculty. In this case we need to think the relation of the agent with himself under the *division* of passivity and activity above sketched. Hence, some contents stand in need of interaction with other persons to acquire meaning, while others are the results from solitary reflection and deliberation.

5. Final Remarks

Here I would like to comment and some of the ideas we have been pursuing throughout this chapter. I'll start by summarizing the two main positions we have presented until now and then I'll criticize Korsgaard's solution to Parafit's example making use of Cavell's concept of rationality and conversation. This criticism aims at a different concept of recognition and its relations to ethics.

So, first I recount our former results. I started with Nagel, whose deep analysis of the debate between consequentialists and deontologists led us to accept the distinction between an agent-relative and an agent-independent moral point of view. This distinction is one of the threads that connects the whole chapter in our quest to understand what is one's singularity in contrast to one's universality. Agent-relative perspective can be a source of duties and moral

dilemmas that stands in contradiction with a agent-neutral perspective. Nagel gave us many examples where one wrongs another person even acknowledging her. This is possible because the moral context made the capacity to act on the basis of agent-neutral duties impossible, but the recognition of the other's humanity was somehow present. Nagel's most remarkable example is the justification a soldier gives to his enemy when killing him. Killing someone during wartimes still is a morally wrong action, according to Nagel, but it can be humanized if both enemies forgive each other because of their awareness that this situation is only lived by them because of the horrible world they live in where many relations do not dependent on their will.

The next step was dedicated to the investigation of the sources of moral normativity present in act of acknowledgement. We followed Darwall's arguments concerning the second person standpoint. He claims that second person competence has the form of moral obligation, for our moral normativity is rooted in the practice of mutual accountability, of reason giving and in our capacity to demand authority and responsibility as member of a moral community. His derivation of the principle follows the clues left by Strawson, where reactive attitudes are the first sign of our attribution of moral responsibility.

Finally, we followed Korsgaard's criticism of Darwall; in her opinion no moral necessity follows from the capacity of having a second person standpoint. She believes the capacity of having a second person perspective on oneself comes from our self-conscious nature. Korsgaard's project is grounded in two points: first our self-conscious nature, and second, her understanding of action as self-constitution. She believes the agent constitutes his or her unity through action, but in order to act morally one must address oneself. In self-address I discover morality in the form of obligations that I have to myself. Self-address is how I interact with myself and in this interaction I must constitute a unity -- a morally integral subject.

Both Darwall and Korsgaard claim morality to be fundamentally dependent on address. Darwall believes we can derive the form of morality from address; Korsgaard believes address and morality are tied to self-consciousness. However, address only makes sense if we are to acknowledge the existence of an addressee, therefore both authors must agree that recognition is tied to morality and address. Nonetheless, how does recognition relate to both concepts?

A good answer to this comes from Cavell's *The Claim of Reason*. In considering Stevenson's argument in *Ethics and Language* about the nature of morality, Cavell tries to show that "he is denying [...] that there are any methods of supporting ethical conclusions which we can sensibly be said to recognize as valid [...]." (p.277). Hence, for Stevenson, contrary to scientific and logical judgments, ethical arguments are just "propaganda" (p.288), destituted of validity. Stevenson says that

"Any statement about any matter of fact which any speaker considers likely to alter attitudes may be adduced as a reason for or against an ethical judgment (Stevenson in Cavell 1979. p. 114).

Cavell reads this quotation such that, if Stevenson is right, then any reasons strong enough to change attitudes can be adduced as a moral justification. It is as if we did not need training and education in moral matters, because as long as you are capable of giving reasons, any statement about any matter of fact can play the role of a reason in a moral debate. Hence, it is as if there weren't specific rules and facts that concern morality specifically, for any reason you give in a moral debate can be adduced as a legitimate reason. However, if any reason is a reason in moral debate, therefore there are no rules to establish legitimacy of arguments.

I understand the denial of validity of ethical arguments as a form of moral skepticism, but instead of immediately saying that there are no objective moral distinctions, Stevenson simply puts the burden on the way we play the language game of moral justification. He considers the practice of moral justification so open that it becomes a groundless practice. Cavell in his answer to Stevenson affirms this conclusion is frustrating in many levels. First because it fails in preserving something as basic as moral concept's functionality for beings like us, but secondly -- and more important -- it is frustrating because it fails to meet others.

As Wittgenstein shows us at the opening of the *Blue Book*, some of the most bewildering but also confusing arguments in Philosophy come from false analogies. Each language game has its own rules, for although they obviously have resemblances, they also have fundamental differences. As Cavell says comparing scientific and moral language games, each language game expresses a commitment to certain modes of argument:

If what makes science rational is not the fact of agreement about particular propositions itself, or about the acknowledged modes of arriving at it, but the fact of a *commitment* to certain modes of argument whose very nature is to lead to such agreement, then morality may be rational on exactly the same ground, namely that we commit ourselves to *certain* modes of argument, but now ones which do not lead, in the same ways, and sometimes not at all, to agreement (about a conclusion).

Thereby we don't need to arrive always at the same moral conclusions in order to agree -- we don't need to agree about what is the exact formulation of the categorical imperative --, but we have to agree in terms of *how* we play the moral language game -- what are the moral notions about, its particular modes of inferences, what are the objects in question, etc.

So, before considering Korsgaard's and Darwall's disagreement, skepticism about morality show us that we should step back from the debate and, instead of taking sides, see how we begin to play the language game of morality. Here Cavell has something important to teach us.

If skepticism about morality entails the lack of recognition of the other, then morality is a problem fundamentally connected to recognition (acknowledgement in the Cavellian sense). Now, if we go back to Cavell's teachings about the concept of acknowledgment, in *Knowing and Acknowledging* we see, as I said in chapter one, that Cavell claims acknowledgment to be the expression of one's knowledge of the other. From these two statements, we should accept, then, that acknowledgment demands a *moral expression of my knowledge of the other*.

If my attitude towards him expresses my knowledge that he has a soul, my attitude may nevertheless not be very definitely expressed, nor very readily. It may take ages; it may be expressed now in the way I live. [...] The word "attitude" can be misleading here. It is not, in the matters at hand, a disposition I can adopt at will. It is helpful to take the English word in its physical sense, as an inflection of myself toward others, an orientation which affects everything and which I may or may not be interested in discovering about myself. (Cavell. 1969. 360)

Cavell uses the word "attitude" to talk about the expression of my knowledge of the other. The word comes from Wittgenstein, who speaks about attitude in what was called Part II of the *Philosophical Investigations*. He uses the word in different connections (§284; §310; §575; §672-3 Part II, iv, §22; Part II, xi, §129, etc) but the sense of "inner change" is the one that he insists in exploring. Attitude is not only a reaction, rather it is what can provoke a reaction, it is a change in the way we see things. The best example is also the most beautiful:

My attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul. (PI, Part II, iv, §22)

Notice that Wittgenstein talks about an attitude *towards a soul*, as if in the quality of being *toward* suggested some directedness, as address was first defined by Darwall. That's why Cavell understands that term as saying that this is not the same as a disposition one can adopt at will. "Attitude" here is an "inflection of oneself towards others, an orientation that affects everything" (CoR, 360). Cavell's understanding is that an attitude opens meaningful ways of

existence. Or less poetically, an attitude enables one to meaningfully play certain language-games that otherwise would remain covered to us.

Wittgenstein insists on the point of attitude. I don't think it is necessary to go deep on this matter here⁴², since my aim in these final remarks is to show how the concept of attitude is connected to morality, because attitude is a note of the concept of address. But I do want to emphasize a further point that I will explain briefly in what follows, namely that according to this understanding an attitude is what one needs to change in order to avoid the phenomenon (what Cavell calls) of aspect-blindness. Aspect blindness occurs when a person lacks the ability to see something *as something*. The classic case of the Duck-Rabbit, shows us two aspects in the same figure. The duck-aspect does not cover the rabbit-aspect, since they both are always there, simultaneously in the picture. What covers or hides aspects are us. Why us? What we can see is more than merely what the eyes passively get from the environment. The eyes are active in connecting aspects, in enabling the perception of relations. Hence, what we see is also what we learn to see, what we have the ability to connect. Hence, it's up to us to fail or succeed in altering our own perception by *continuously connecting aspects* so as to see the duck-aspect instead of the rabbit-aspect, this person is said to be aspect blind. (PI, Part II, xi, §257)

Aspect blindness is opposed to aspect dawning. An aspect dawns on us when it is revealed or uncovered to us. If one fails to connect aspects continuously so as to see a physiognomy, then this person remains aspect-blind. The way Wittgenstein describes this happening is through a change in the way one continuously connects aspects. As when one connects the ears of the rabbit in connection to the nose, but suddenly one connects the same parts of the picture differently, say the duck's beak with the top of its head. This change in perception has an analogous case in terms of attitude. Attitudes reveal things to us, as Cavell describes

If I say, for example, that I do not mind the cold and offer you my jacket, does this necessarily mean that I am not as cold as you? [...] To me this seems an image of freedom. It seems to me also an example of the possession, or exercise, of a will. But here, I find, I am not thinking of the will as a kind of strength which I may have more or less of, but as a perspective which I may or may not be able to take upon myself. So one may say that the will is not a phenomenon but an attitude toward phenomena. (Cavell. 1979, p.361)

⁴² At least two great books that do that come to my mind in this connection. *Mulhall, Stephen (1990). On Being in the World: Wittgenstein and Heidegger on Seeing Aspects. And Day, William & Krebs, Victor J. (eds.) (2010). Seeing Wittgenstein Anew.*

So an attitude enables one to get in touch with perspectives "I may or may not be able to take upon myself". The same is said in the first chapter, when Cavell describes the difference between the situation in which of someone (merely) *knows* he or she is late and the situation in which one *acknowledges* that one is late. If I know I am late, it does not mean I do something about it. Knowledge does not entails a change of attitude. Some people know they are sick and, even though, they don't look for treatment. They don't acknowledge that this a form of degradation probably motivated by unconfessed reasons. Another example: If I were a teacher that does not care about my students, my belatedness would make no changes on my attitude. It is simply knowledge. However, if I acknowledge I am late, I change my attitude in this respect. Perhaps I'll try to hurry up, given my sense of responsibility; perhaps I'll simply realize that I'm late because I'm not happy with my job. What is important here is that acknowledgment leads to a change of attitude, for it leads to a change of perspective one takes upon oneself. This capacity to change one's perspective on oneself is what I take to be at the basis of what connects a person to morality.

Returning to Cavell's criticism of Stevenson, what I think is implicated in the way Cavell understands skepticism about ethics is that moral justification is a practice that depends on an attitude towards oneself and towards others. Moral justification is a language game that one plays both with oneself and with others. This is what enables one to recognize oneself as an integral and moral person, but also this is what makes possible that others recognize this very important social feature in one. So, both authors, Korsgaard and Darwall agree that address is the engine of morality, but also both agree that address is only possible with acknowledgement. She believes morality comes from the capacity to address myself, while Darwall believes morality comes from the capacity to address others. What we learn from Wittgenstein's idea of aspect-blindness and aspect downing as ways to see physiognomies is that we don't need to choose between Korsgaard and Darwall. Context will be more important. Someone who's afraid of looking inwardly will be blind to the morality manifested in the way one treats oneself. So, this person will discover morality through the presence of others in her life. Someone who's afraid of the other, will learn the grammar of morality interacting with oneself. What determines one or the other? I like to think that the conditions of human life is what determines which kinds of address one will take as reliable. Perhaps you were still young when war and crime were epidemic in your country, so trust and confidence in others were not easy. Perhaps you were raised in a rich family that gave you all the attention and care needed for you to believe other's words.

Morality emerges from an attitude contrary to Stevenson's skepticism. This attitude enables one to see the good and the evil, to see each other's identities, to act, to refrain from act and all sorts of concepts and judgments we use to talk about this aspect of our form of life.

Once one steps in this region of our form of life, a difference can be settled: One thing is to fail to justify one's own actions to oneself, other totally different thing is to unknowingly cause horror and pain because one is failing to acknowledge morality as a facet of the human condition. Morality depends on being able to see oneself and others as persons. This uptake on others is the primary openness to morality -- I would risk to say, the primary source of what we want categorical imperative to be, if it is to be the form of human concern with the other. This is (also) what the debate about seeing aspects and seeing the physiognomy of persons is about.

The block to my vision of the other is not the other's body but my incapacity or unwillingness to interpret or to judge it accurately, to draw the right connections. The suggestion is: I suffer a kind of blindness, but I avoid the issue by projecting this darkness upon the other. The convincingness of Wittgenstein's thought here will depend upon whether one is convinced by his relating the condition of what he calls "seeing an aspect" (hence what he calls "aspect-blindness") to seeing a likeness between physiognomies, which he cites as his motive in introducing the topic of aspects in the first place. Aspect-blindness is something in me failing to dawn. It is a fixation. In terms of the myth of reading the physiognomy, this would be thought of as a kind of illiteracy; a lack of education. (Cavell, 1979, 368-9)

I take this to be in connection to his examination of Stevenson's arguments. Cavell's understanding of aspect blindness takes this phenomenon to be a "kind of illiteracy; a lack of education". Therefore, as morality is connected to the capacity to see someone *as a person*, it is connected to the practice of acknowledgment.

To finish this little and still to be developed (in future articles) excursus on Cavell's understanding of morality, I would like to get back to Korsgaard's noble socialist. This example is the classic case of keeping one's word:

The story goes like this. The nineteenth-century Russian is now, in his youth, a socialist, and he plans to distribute large portions of his inheritance, later, when he comes into it, to the peasants. But he also anticipates that his attitudes will become more conservative as he grows older, and that he may not think that this is the right thing to do when the inheritance is finally his own. So he makes a contract *now*, to distribute the land when he gets it, which can only be revoked with the consent of his wife, and he asks his wife to promise not to revoke it *then*, even if he

tells her *then* that he has changed his mind, and that she is released from the promise. (Korsgaard, 2009, p.185)

Korsgaard analyses this example to comment how one ought to interact with oneself. She claims integrity to be the goal of an agent with the capacity to be an end-in-itself. Integrity is the capacity to constitute a unified agent where motivations, actions and ends are all integrated. The socialist nobleman example allows Korsgaard to show us what is an agent with no unity -- lacking integrity. Hence, an agent whose motivations are not followed in his actions throughout time.

The example reaches its goal in one aspect: in defending a concept of integrity as a form of coherence with oneself. However the picture of "self" and the metaphor of unified identities as parts of the self, hardly escapes to a form of temptation. The socialist asks his wife to not allow him to betray his own words, hence the promise is also made by his wife, since she is the person in charge and responsible for preventing him of doing things otherwise. But the example puts as the only condition for the promise to be kept another will. The will of the nobleman's wife. Integrity can't be a mode of coherence with oneself if the self is to be thought as detached from the world in a past time. The self is connected to the world and it can reasonably change its beliefs and world-view. Old promises and objectives lose vitality, the life questions our promises, they can lose meaning because the world demands inner changes. I don't want to believe that the jew father who was living a decent life as a barber in Paris, safe from problems until the Nazis took the government, lost his integrity because he had to abandon one of his identities. His life changed priorities from being a good barber to keep his son safe and it seems all right to say that integrity can't be a simple sum of identities fully constituted with the help of virtue; it seems that abandoning some identities is a fundamental step in rationality, specially when the world, your country and your society has no longer a place for it.

Cavell in his analysis of the Comedies of Remarriage seeks to show us a picture of rationality connected to the low, the familiar and the ordinary. He claims that a set of related films from 1933 to 1949 compose a genre of comedies where the dialogue between a man and a woman is the development of a relation of recognition.

Our films may be understood as parables of a phase of the development of consciousness at which the struggle is for the reciprocity or equality of consciousness between a woman and a man, a study of the conditions under which this fight for recognition (as Hegel put it) or demand for acknowledgement (as I have put it) is a struggle for mutual freedom, especially of the view each hold of the other. (Cavell.1981 , p. 17-8)

These comedies can help us here to address the example of the socialist nobleman. Remarriage is possible in these comedies because the couple goes through a process of recounting of the ordinary. The affirmation of love comes not from a big event, as in the the Shakespearean comedy, but from a return to the basic sharedness of life.

Sharing an ordinary life, the low and the common is a way to get in touch with some fundamental form of relatedness to the world and to others. Cavell's reading of Wittgenstein's *Philosophical Investigations* enables an understanding of modern skepticism as an expression of intellectual temptation where agreement in criteria is lost. The need for certainty present in modern skepticism leads the skeptic to no longer believe he can know about the existence of any objects. If he can't know about the existence of an apple, simply because he can't completely see the apple in front of his eyes (given that there's always an disclosed side), this means that even a *best case of knowledge* does not suffice to obtain knowledge. The skeptic forgets his finitude when asking for a knowledge out of perspective.

Mulhall, one of the most brilliant philosophers to follow Cavell's lessons, tells us that the when the skeptic tries to have knowledge out of a finite perspective, he forgets the grammar of our language.

The grammar of our language determines a logical space of possibilities, and the world determines which of these spaces are filled; we can investigate the latter only by looking to see what is the case, but we can investigate the former from our armchair by utilizing our knowledge of the criteria governing the application of the words of our natural language to the world. (Mulhall, Stephen. 1994 p.18)

So, the grammar of an object (as a chair is to sit, to block a door, to hit someone, etc) is governed by criteria. If criteria tells us what is an x in reality; what kind of object something is -- a material object, a person, a divinity --, then criteria is at the basis of the experience of reality, it is the expression of our form of life⁴³; of how we relate to this world and which world one is to find when one shares criteria in a community.

Now back to the the socialist nobleman, he should have nourished his promise during his life. Throughout everyday life he could have acted according to the purpose of fulfilling his promise. If he failed to continue to be the same person during his lifetime, this is probably

⁴³ Cavell insists on the possibility of variation. The concept of form of life has a vertical variation, where we are taken as humans compared to other species. In this perspective we can see biological similarities with other species and differences. The concept of form of life has also the horizontal variation, where we compare humans among their different cultures. This is the anthropological sense of the term.

because he constituted his everyday life in a different direction from what he has imagined when he was young. But we don't know what was his world. I insist with barber's example: coherence with oneself depends of the world as well.

Now he needs his wife to the keep his word, but she's no longer talking to that socialist young man. Korsgaard says she must treat him as a disunified person. I think that, analogously to what happens with Wittgenstein's stubborn math student: instead of adopting Kripke's solution to the paradox of rule following, where he suggests we should admit at some moment that we "act blindly", Cavell's solution points to the possibility of conversation. Cavell suggests that remarriage depends on the possibility of sharedness, of conversation. Conversation is the path to recount criteria, to reach back the basic sharedness of life. So, the noble socialist's wife could try to find again the young socialist in her husband and remarry him. If the comedies can teach us something, their teaching is that talking is the way rationality is delineated, for rationality is rooted (*ratio*) in everyday life, in the life of beings whose burden is to think and talk in order to solve their problems and satisfy necessities from both the stomach and the heart. Perhaps his commitment to his past self won't make sense any longer, and our conversation (our rationality) should allow us to understand that. Perhaps recounting of the ordinary through conversation they can find in the old nobleman's actual self the scattered pieces of a young socialist. Integrity is not to be found merely in coherence, but in the capacity for rational transfiguration.

Conclusão

Ao longo desta tese uma diversidade de assuntos e de autores foram abordados, e talvez não tenha sido sempre claro qual é o princípio de unidade dessas análises. Expressar essa unidade que ao longo do caminho creio ter delineado será a tarefa desta breve conclusão.

Primeiramente, esta tese busca expressar o resultado de uma experiência um tanto pessoal, mas certamente aberta a todos aqueles que buscam estudar o ceticismo. Aprendi que o ceticismo não é meramente um problema restrito ao âmbito filosófico. O problema cético se manifesta de inúmeras formas e onde quer que se possa pedir aos homens e mulheres que sigam sendo racionais. Seja na literatura, no cinema, no teatro; seja na ciência, na epistemologia, na ética; seja no casamento, no amor ou amizade; em qualquer uma dessas esferas o ceticismo pode vir a assombrar a relação com seu objeto. Ele é a expressão de algo que podemos chamar de "quebra" ou de "ruptura" do self nas suas relações mais fundamentais com o mundo e com os outros.

Em segundo lugar, mas igualmente importante, minha busca ao longo da tese foi responder à questão: o que é reconhecer o outro? Iniciei a análise através da conexão que me levou até ela: mostrando que o diagnóstico do problema cético sobre outras mentes nos leva a compreender este antigo desafio epistemológico como um algo ligado ao reconhecimento do outro. A partir do momento que pude fazer essa conexão, minha busca passou a ser analisar o fenômeno (e o conceito) de reconhecimento que aparece em diferentes tradições filosóficas e em diferentes usos.

Acredito que a primeira conclusão que obtive foi perceber como a estratégia argumentativa que busca conhecer outras mentes por analogia é insuficiente. Diante da insuficiência de tal resposta o problema nos levou a adotar uma postura diferente, uma postura na qual a investigação acerca do conceito de mente compromete a capacidade de fazer sentido de si mesmo, do que é um corpo expressivo e do que é "uma mente". Uma tal incursão -- que parte de uma questão sobre o conhecimento de outras mentes e vai às condições de possibilidade desse conhecimento -- mostrou uma forma de conexão entre as variações cartesiana e kantiana de ceticismo.

Assim, precisamente na medida que o problema cético deixa de ser um mero puzzle filosófico, ele passa gradualmente a ser uma experiência de auto-conhecimento. Esse autoconhecimento progride na medida que a conexão entre ceticismo e reconhecimento fica mais clara, pois as extravagantes exigências céticas só são colocadas de lado na medida que o reconhecimento do outro (e conseqüentemente de si) permite ver a própria finitude.

Quanto aos quatro problemas anunciados na introdução, acredito ter em alguma medida dado uma resposta a eles.

O primeiro problema tratava da unidade do fenômeno do reconhecimento. Afinal, o que compõe o fenômeno do reconhecimento quando alguém reconhece o outro? Ao longo da tese vimos que há formas do reconhecimento que não são mútuas. São formas do reconhecimento que se dão em um sentido só. No entanto, onde quer que haja reconhecimento do outro (seja mútuo ou unilateral), haverá identificação de alguém enquanto pessoa. O conceito de pessoa qualifica alguém como dotado de racionalidade prática e validade social, como argumentei. Todavia, esse sentido preciso de reconhecimento mobiliza outros dois: o sentido de re-identificação de algo previamente identificado, e o sentido de admitir algo para si mesmo. Neste caso, o que é re-identificado e admitido é o senso próprio de pessoa que carregamos conosco. Por isso é comum dizer que o reconhecimento do outro conduz a uma experiência de encontro de si mesmo no outro. O outro mobiliza uma espécie de reencontro consigo mesmo através do que foi chamado de percepção judicativa -- o modo próprio como percebemos o outro mobiliza uma percepção de si mesmo.

Não obstante, há outras formas de reconhecimento que não são unilaterais, mas que são chamadas de mútuas. O que nos leva a concluir que há diferentes maneiras pelas quais o fenômeno adquire unidade. No caso reconhecimento mútuo vimos que há um debate sobre como isso ocorre. Vimos que para compreender a lógica desse fenômeno não cabe simplesmente pensar que o pronome pessoal de primeira pessoa do singular mais o demonstrativo "este" referindo ao outro. Essa forma de entendimento isola os polos da relação em pilhas de pensamento que não tem qualquer garantia de estarem engajados na relação. Uma tal explicação do fenômeno trata as pessoas como ilhas que se referem umas às outras sem que haja reconhecimento propriamente. O saldo desse debate mostrou que para pensar genuinamente o conceito de reconhecimento é necessário tomar essa relação como compondo uma unidade referida pelo pronome pessoal de primeira pessoa do plural. Tal pronome é pensado em fenômenos de agência mútua (ou compartilhada), triangulação e atenção conjunta.

Outro problema que me interessou ao longo da tese foi a capacidade de reconhecer o outro na sua diferença mais radical. Busquei explorar uma da forma de manifestação do ceticismo sobre outras mentes chamada de cegueira para o outro. A cegueira para o outro se manifesta na forma de invalidação social do outro, assim como pela projeção de crenças e atitudes sobre o outro. Para acessar o outro na sua alteridade é necessário que partes pouco exploradas da nossa humanidade estejam acessíveis a nós, é necessário um exercício de auto-endereçamento (self-address) onde identidades possíveis sejam significadas, onde buscamos iluminar as partes que são obscurecidas por nós mesmos. Pareceu-me que através

desse exercício é possível conectar-se com o outro "lá" onde ele espera/deseja ser endereçado.

O endereçamento ao outro na sua alteridade é condição para que o pronome pessoal de primeira pessoal do plural seja representativo das formas do reconhecimento em situações de agência mútua. O endereçar-se é o modo próprio pelo qual o outro é abordado nas suas identidades pessoais. Tais identidades e sua conexão com o tema do reconhecimento das outras vontades que não a minha própria foram tema do meu interesse no último capítulo da tese.

Finalmente, como anunciei na introdução, meu interesse final seria conectar o tema do ceticismo sobre outras mentes com o tema do reconhecimento de outras vontades. Reconhecer o outro no plano ético é tratá-lo como um sujeito capaz de agência. Enquanto tal, o que se reconhece é um sujeito dotado de racionalidade prática e de algum senso de responsabilidade. A implicação imediata desse reconhecimento é moldar a forma que minha vontade deve ter para que ela respeite o outro enquanto um ser livre. O modo como posso moldar minha vontade ao outro é através do reconhecimento do outro num nível pessoal. Reconhecer o outro demanda saber quem ele é, permitir que ele se abra a mim. Isso se dá, novamente, através do exercício de endereçamento, e neste caso, através da prática de justificação moral. Vimos que dois autores divergem acerca do que é fundamental em tal prática, se o outro hipotético ou o outro que somos nós mesmos enquanto seres auto-conscientes. Sobre isso, preferi não tomar partido. Acredito que o jogo de linguagem onde pesam os termos morais são jogados em ambas esferas sem que haja preponderância de uma sobre a outra. São conversas distintas e elaborações de formas da racionalidade de alcance distinto - uma tendo a si mesmo como fim e outra tendo outras como fim.

Há uma última conclusão que gostaria de extrair aqui, para a qual certamente não dei suporte teórico sistemático, mas que pode ser entrevisto nas inúmeras referências que faço à busca por autoconhecimento que o ceticismo propicia. Minha jornada na filosofia se iniciou ao entrar em uma sala do curso de filosofia da UFRGS em março de 2009. Nessa aula escutei por primeira vez a máxima aristotélica de *Metafísica Alfa*: "Todo homem deseja por natureza conhecer". Nove anos depois o mal-estar ocasionado por essa frase finalmente foi colocado para descansar. Parece-me -- e escrever esta tese reforçou essa fortíssima impressão -- que o que ao menos alguns de nós desejam por natureza é reconhecer e ser reconhecido.

Bibliography

- Affeldt, Steven G; ", "The Ground of Mutuality: Criteria, Judgment and Intelligibility in Stephen Mulhall and Stanley Cavell", *European Journal of Philosophy*, 6, 1, 1-31, 1998, Wiley Online Library
- Austin, John Langshaw. *How to do things with words*. Oxford university press, 1975.
- Avramides, Anita. *Other minds*. Routledge, 2000.
- Ayer, A.J. *The Foundations of Empirical Knowledge*. (1940)
- Cavell, Stanley; ", "The claim of reason: Wittgenstein, skepticism, morality, and tragedy" ,,,,,, 1999, Oxford University Press
- Cavell, Stanley; ", *Must we mean what we say?: A book of essays*, ,,,,,, 2015, Cambridge University Press
- Cavell, Stanley; ", *Disowning Knowledge: In Seven Plays of Shakespeare*, ,,,,,, 2003, Cambridge University Press
- Cavell, Stanley; ", *The world viewed: Reflections on the ontology of film*, ,,,,,, 1979, Harvard University Press
- Cavell, Stanley; "Pursuit of Happiness." *The Hollywood Comedy of* (1981).
- Cavell, Stanley; ", *Knowing and acknowledging, Must we mean what we say*, 238,,266, 1969,
- Child, William; ", *Remembering intentions*, ,,,,,, 2010, na
- Churchland, Paul M. *Scientific realism and the plasticity of mind*. Cambridge University Press, 1986.
- Day, William & Krebs, Victor J. (eds.) (2010). *Seeing Wittgenstein Anew*. Cambridge University Press..
- Davidson, Donald. "Rational animals." *dialectica* 36.4 (1982): 317-327.
- Descartes, R; ", *"Meditations on First Philosophy. VI, 80//Descartes. Discourse on Method and Meditations on First Philosophy/trans. DA Cress"*, ,,,,,, 1998, Indianapolis/Cambridge: Hackett Publishing Company
- Darwall, Stephen; ", *III—Moral Obligation: Form and Substance, Proceedings of the Aristotelian Society (Hardback)*, 110, 1pt1, 31-46, 2010, Wiley Online Library
- Darwall, Stephen L; ", *"The second-person standpoint: Morality, respect, and accountability"*, ,,,,,, 2006, Harvard University Press
- Diamond, Cora; ", *"Realism and the Realistic Spirit: Wittgenstein, Philosophy, and the Mind"*, ,,,,,, 1991, "Cambridge, Mass.: MIT Press"
- Evans, Gareth; McDowell, John Henry; McDowell, John Henry; ", *The varieties of reference*, ,,,,,, 1982, Clarendon Press Oxford

- Finkelstein, David H. Expression and the inner. Harvard University Press - 2008
- Foot, Philippa; ",Natural goodness,,,,,2003,Clarendon Press
- Frege, Gottlob, and Michael Beaney. "The Frege Reader." (1997)
- Gallese, Vittorio. "The 'shared manifold' hypothesis. From mirror neurons to empathy." Journal of consciousness studies 8.5-6 (2001): 33-50.
- Gallagher, Shaun & Varga, Somogy (2014). Social Constraints on the Direct Perception of Emotions and Intentions. Topoi 33 (1):185-199.
- Gallup Jr, Gordon G; Anderson, James R; Shillito, Daniel J; ",The mirror test,The cognitive animal: Empirical and theoretical perspectives on animal cognition,,,325-333,2002,"MIT Press, Cambridge, MA"
- Geach, Peter T; ",On beliefs about oneself,Analysis,18,1,23-24,1957,Wiley-Blackwell
- Geach, Peter T; ",Good and evil,Analysis,17,2,33-42,1956,JSTOR
- George, Alexander. "Reflections on Chomsky." (1989).
- Gobel, Matthias S., Heejung S. Kim, and Daniel C. Richardson. "The dual function of social gaze." Cognition 136 (2015): 359-364.
- Goldfarb, Warren; ",Wittgenstein on understanding, Midwest studies in 17,1,109-122,1992,Wiley Online Library
- Haddock A (2014) On Address, Philosophical Topics, 42 (1), pp. 345-350.
- Hume, David; ",A treatise of human nature,,,,,2003,Courier Corporation
- Honneth, Axel, and Avishai Margalit. "Recognition." Proceedings of the Aristotelian society, supplementary volumes 75 (2001): 111-139.
- Kant, Metaphysics of Morals (1797) Introduction, Translation and notes by Mary Gregor. Cambridge University Press (1991)
- Kripke, Saul; ",Wittgenstein on Rules and Private Language,,,,,1982,
- Kobayashi, Hiromi, and Shiro Kohshima. "Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye." Journal of human evolution 40.5 (2001): 419-435.
- Korsgaard, Christine M; ",,"Self-constitution: Agency, identity, and integrity",,,,,2009,OUP Oxford
- Korsgaard, Christine; ",Aristotle's function argument,,,,,2008,Oxford University Press
- Korsgaard, Christine M; ",Autonomy and the second person within: A commentary on Stephen Darwall's The Second-Person standpoint, Ethics,118,1,8-23,2007,The University of Chicago Press

- Malcolm, Norman; ",Knowledge of other minds,Essays in philosophical psychology,,
365-376,1964,Springer
- Margalit, Avishai; ",The decent society,,,,,1998,Harvard University Press philosophy, Studies,
- Margalit, Avishai; ",Recognition: Recognizing the brother and the other: Avishai
Margalit,Aristotelian Society Supplementary Volume,75,1,127-139,2001,Wiley Online
Library
- McDowell, John; ",Referring to oneself,The Philosophy of PF Strawson,,,127-146,1998,
- McDowell, John; ", "Criteria, defeasibility, and knowledge" ,,,,,1982,
- McDowell, John; ",One strand in the private language argument,Grazer Philosophische
Studien, 33,,285-303,1989,
- McDowell, John; ",Wittgenstein on following a rule,Synthese,58,3,325-363,1984,Springer
- McDowell, John Henry; ", "Mind, value, and reality" ,,,,,1998,Harvard University Press
- McDowell, John; ",Two sorts of naturalism,"Mind, value, and reality",167,,97,1998,Cambridge:
Harvard University Press
- Mill, John Stuart. An Examination of Sir William Hamilton's Philosophy: And of the Principal
Philosophical Questions Discussed in His Writings. Longmans, Green, and Company,
1889.
- Moran, Richard; ",Cavell on outsiders and others,Revue internationale de philosophie,,
2,239-254,2011,Association Revue internationale de philosophie
- Mulhall, Stephen; ",On Being in the World (Routledge Revivals): Wittgenstein and Heidegger
on Seeing Aspects,,,,,2014,Routledge
- Mulhall, Stephen. Stanley Cavell: Philosophy's Recounting of the Ordinary. Clarendon Press,
1994.
- Mulhall, Stephen; ",The Cavell Reader,,,,,1996,
- Mulhall, Stephen; ",Seeing aspects,,,,,2001,
- Nagel, Thomas; ",The view from nowhere,,,,,1989,oxford university press
- Nagel, Thomas; ",The objective self,Knowledge and Mind,,,211-32,1983,
- Nagel, Thomas; ",War and massacre,Philosophy & Public Affairs,,,123-144,1972,JSTOR
- Onishi KH, Baillargeon R. Do 15-Month-Old Infants Understand False Beliefs? Science (New
York, N.y
- Peacocke, Christopher, "The mirror of the world: subjects, consciousness, and self-
consciousness" ,,,,,2014,Oxford University Press
- Parfit, Derek; ",Reasons and persons,,,,,1984,OUP Oxford

- Perry, John; ",The problem of the essential indexical,Noûs,,,3-21,1979,JSTOR
- Putnam, Hilary. "Brains and behavior." Readings in philosophy of psychology 1 (1980): 24-36.
- Recanati, François; ",Immunity to error through misidentification: what it is and where it comes from,,,,, 2012,Cambridge University Press.
- Reid, Thomas; ",The works of thomas reid,,1,,1895,James Thin
- Rödl, Sebastian; ",Self-consciousness,,,,,2007,Harvard University Press
- Rödl, Sebastian; "Intentional transaction, Philosophical explorations,17,3,304-316,2014, Routledge
- Russell, Bertrand; ",Analogy,Essays on other minds,,,3-8,1970,
- Russell, Bertrand. "On denoting." Mind 14, no. 56 (1905): 479-493.
- Russell, Bertrand. The problems of philosophy. OUP Oxford, 2001.
- Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." Minnesota studies in the philosophy of science 1.19 (1956): 253-329
- Strawson, Peter Frederick. Individuals. Routledge, 2002.
- Tomasello, Michael. "Joint attention as social cognition." Joint attention: Its origins and role in development 103130 (1995).
- Wittgenstein, Ludwig; ", "Philosophical investigations (GEM Anscombe, trans.)",,,,,1953,Oxford: Blackwell
- Wittgenstein: Tractatus Logico-Philosophicus, Philosophy of Language: The Key Thinkers 76,2011,A&C Black
- Wittgenstein, Ludwig; ",Zettel,,,,,1967, University of California Press