Tesis Doctoral

Talking the Way to Other Minds: Assessment, Conversation and Folk psychology

Doctorando

Víctor Fernández Castro

Directores

Dr. Fernando Martínez Manrique

Dr. Agustín Vicente Benito

Dpto. de Filosofía I

Universidad de Granada

Editor: Universidad de Granada. Tesis Doctorales
Autor: Víctor Fernández Castro
ISBN: 978-84-9125-140-8
URI: http://hdl.handle.net/10481/45377

El doctorando / The *doctoral candidate* **[Víctor Fernández Castro]** y los directores de la tesis / and the thesis supervisor/s: **[Fernando Martínez Manrique y Agustín Vicente Benito]**

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

/

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / Place and date:

Granada, 25 de Enero de 2017

Director/es de la Tesis / *Thesis supervisor/s;* Doctorando / *Doctoral candidate*:

Firma / Signed

Firma / Signed

A MI LOLA

Publications

Víctor Fernández Castro (forthcoming) "Inner Speech in Action" *Pragmatics & Cognition* 23(2).

Víctor Fernández Castro (forthcoming) "Mindshaping and Robotics" in Hakli & Seibt (eds.) *Sociality and Normativity for Robotics: Philosophical Inquiries into Human-Robotics-Interactions*, Springer.

Víctor Fernández Castro (2016) "Narrativity and Propositional Attitudes: Beyond an Expressivist View of Folk Psychology" in Sociedad Española de Filosofía Analítica (eds) *Eight Meeting of the Spanish Society for Analytic Philosophy.* KRK Ediciones, 56-58.

Víctor Fernández Castro (2015) "The Normative Dimension of Folk Psychology" in Díez, García-Carpintero, Martínez y Oms (eds.) *VIII Conference of the Spanish Society for Logic, Methodology and Philosophy of Science.* University of Barcelona, 15-16.

Víctor Fernández Castro (2015) "Dar Forma a la Mente" *Thémata. Revista de Filosofía* 51, 433-442.

Víctor Fernández Castro (2015) Review: Tadeusz Zawidzki (2013) Mindshaping, Cambridge: MIT Press. *Teorema* 34(1): 187-191.

Víctor Fernández Castro (2014) "Shaping Robotic Minds" in Seibt, Hakli & Nørskov (Eds.) S*ociable Robots and the Future of Social Relations*, IOS Press (71-78).

Víctor Fernández Castro (2012) "Metarepresentación y Dualismo en la Lectura de mentes" In Martínez, Falguera, Sagüillo, Verdejo & Pereira-Fariña (Eds.) *VII Conference of the Spanish Society for Logic, Methodology and Philosophy of Science.* Servicio de Publicación de la Universidad de Santiago de Compostela, 184-190.

Víctor Fernández Castro (2012) Review: Antoni Gomila (2012) Verbal Mind: Language and the Architecture of Cognition, Amsterdam: Elsevier. *Theoria* 27(3): 394-397.

Agradecimientos

> "The mutual-aid tendency in man has so remote an origin, and is so deeply interwoven with all the past evolution of the human race, that is has been maintained by mankind up to the present time, notwithstanding all vicissitudes of history." — Pyotr Kropotkin

Quizás sean el gen C.N.T.ista de mi familia, o quizás simplemente he tenido la suerte de estar en las circunstancias adecuadas, pero nunca he tenido problemas para darme cuenta, como Kropotkin defiende en esta cita, de que el ser humano tiene una tendencia natural para el apoyo muto. Pero incluso si esto no hubiera sido así, mi propia experiencia vital y estos años que he pasado en el departamento de filosofía I de la Universidad de Granada se hubieran encargado de demostrármelo con creces. Esta tesis doctoral, el trabajo realizado durante estos años, y sobre todo, las cosas que he aprendido, no habrían sido posibles sin una gran cantidad de personas que han puesto su esfuerzo, sus ganas y su cariño en ayudarme para llegar a este punto. Esta tesis, y toda mi gratitud es para ellos.

En primer lugar, me gustaría agradecerle a mis directores Fernando Martínez y Agustín Vicente toda la confianza puesta en mi, las horas de ayuda, los consejos, el tiempo pasado leyendo borradores y papers, las palabras de ánimo, su enorme paciencia y el no haberme dejado nunca solo a lo largo de estos años. Además de su ayuda, ellos han sido siempre un ejemplo a seguir y una fuente de admiración. Hace años, cuando Fernando me inyectó la curiosidad por la

filosofía de la mente y la psicología no sólo puso en mi la semilla de lo que hoy son mis intereses filosóficos, sino que además, siempre ha sido el andamiaje vygotskiano sobre el que apoyarme para desarrollar mis destrezas como filósofo. En sus horas libres como máquina expendedora de ejemplos filosóficos, Fernando me aconsejó trabajar con Gusa durante mi año de master, y aceptar ese consejo ha sido una de las mejores decisiones que he tomado. Durante ese máster, Gusa me introdujo en los problemas centrales de los que se ocupa esta tesis doctoral. Después, y a pesar de la distancia, siempre me ha animado y me ha tendido su mano. Gracias a los dos por enseñarme a ser un mejor investigador, y sobre todo, por vuestra amistad. Gracias de todo corazón.

Neftalí Villanueva y Manuel de Pinedo son dos personas a las que también debo agradecer muchísimas cosas. No sólo por sus aportes a esta tesis, su seguimiento de mi trabajo y su gran influencia en mi filosofía, sino por ser los motores principales de ese ambiente de trabajo perfecto que es el Granada Gang. Asimismo, debo agradecer a Neftalí el haberme introducido en esa secta llamada Brazilian Jiu Jitsu; gracias por esos días de armlocks, estrellas de bodega y conversaciones de rap y filosofía. Al maestro, a quien admiro por sus dotes para consumir cerveza casi tanto como por su agilidad como pensador, tengo que agradecerle haber sido siempre una fuente de conocimiento, tanto en sus clases, al principio de mi carrera, como durante nuestras conversaciones y charlas en todos estos años. Gracias por todo a los dos.

Además de estos dos grandes motores, el Granada Gang tiene los mejores engranajes que una máquina filosófica podría tener: millones de gracias a Tori 'mi cabesa' y Manolo Heras por empezar este camino conmigo (¿Qué sería de nosotros si 'The Lady' no nos hubiera regalado esa taza?). Millones de gracias al dúo dinámico (Nemesio y Jose aka 'Tete Pistolas'), a Eduardo 'Taburete' Pérez, David 'Zaragoza', Andrés, Liñan, Palma, Mirco 'este aka mejor no lo

dejo por escrito', Pedro y Francesco. Sois los mejores compañeros que un filósofo puede tener, y los mejores amigos que un borrachuzo podría soñar. A María José Frápolli tengo que agradecerle enormemente todo el apoyo recibido desde el principio de mi carrera hasta hoy. Si no hubiera sido por sus clases, por sus consejos y por su ayuda, probablemente nunca hubiera llegado hasta aquí. Tengo que dar las gracias también a Juan José Acero por sus increíbles enseñanzas durante estos años, por ese maravilloso curso sobre la filosofía de Wittgenstein y por una gran cantidad de buenos consejos. Agradezco enormemente también el apoyo y la influencia que en mi han tenido otros miembros (actuales, anteriores y futuros) del departamento de filosofía I de la Universidad de Granada, en especial a María José García Encinas, Esther Romero, Cristina Borgoni, Daniel Galdeano, Manu, Llanos, Alba, Xavi y Ricardo.

During these years, I have had the fortune of being awarded with two visiting fellowships, which give me the opportunity of studying abroad at two incredible centers. The first visit was at University of Edinburgh where I could enjoy and take advantage of the valuable comments and conversations with my fantastic supervisor Andy Clark. I am extremely grateful for his advices and the thought-provoking discussions of my work. My second visit was at George Washington University, where I have the opportunity of working with Tad Zawidzki. Tad read carefully many drafts and papers I sent him, and he always sent them back with insightful comments and advices. We had many interesting and productive conversations during this time, and his extraordinary philosophical writings have had an incredible influence on this dissertation. I cannot be more grateful to him. During these visits, I met a huge number of amazing people; but I will never forget my family at the 3719 Georgia Av. Thank you Anna, Hanford, Tran, Lalita, Sudesh, Noël and Horacio. I miss you all.

Gracias a toda mi familia, en especial a mi Primo Alex.

Para mis padres, Conchi y Paco, y para mi hermana Laura, no tengo palabras para agradecerles todo lo que han hecho por mi durante toda mi vida. Gracias por estar siempre ahí, por el amor incondicional, por las palabras de aliento, por creer siempre en mí (y a pesar de todo). Sin vosotros, no sólo no sería un filósofo, sino que ni siquiera sería una persona íntegra. Os Quiero Muchísimo.

Por último, quiero dar las gracias a esa persona que ha sido el centro de mi vida durante los últimos años. Gracias por quererme, por padecerme, por tener siempre una sonrisa y un te quiero para mí. No podría haber hecho nada de esto sin tu amabilidad, sin tu dulzura, sin tu apoyo. Espero que esto sólo haya sido el principio de toda una vida juntos. Te quiero, Sara

Contents

Summary

The main aim of this dissertation is to offer a plausible hypothesis of the relation between language and folk psychology. According to this hypothesis, which I call the evaluative conversational hypothesis, human understanding of other agents in terms of mental states requires mastering certain complex linguistic abilities. In particular, humans need to engage in conversationally mediated joint and cooperative activities in order to acquire the conceptual capacity of ascribing propositional attitudes.

What motivates a philosophical inquiry of the connection between language and folk psychology is the discontent with an assumption shared among the different empirical theories concerning this connection. In particular, these theories assume that propositional attitude ascriptions are descriptions of the inner psychological states of the subject under interpretation. This assumption takes for granted that, as social creatures, humans need to access other agents' internal psychological machinery for the sake of prediction, coordination and explanation. Our ascriptions of desires, beliefs, hopes or other mental states represent or describe those psychological states which bring out courses of behavior that, otherwise, would appear alien to us. This thesis, which I introduce in chapter 3 under the label folk psychological descriptivism, is a heritage of a general approach to language according to which the main function of our expressions is

to describe or denote worldly aspects and objects (Chrisman, 2007).

The theses concerning the relation between language and folk psychology are diverse (Chapter 2), and they oscillate from those which negate any influence of language on social cognition, to those strongly committed to the idea that language is a necessary condition for folk psychological skills (Astington and Baird, 2005). In spite of this diversity, I argue that they all share their commitment to folk psychological descriptivism, and in fact, this commitment plays an important role in their different argumentative strategies (Chapter 3). My central contention is that folk psychological descriptivism is highly problematic. Thus, the descriptivist analysis of propositional attitude ascription must be replaced by an alternative.

In order to make my point, I present three arguments against folk psychological descriptivism (Chapter 4). The first argument presents different everyday uses of propositional attitude ascriptions which are hardly interpretable from a descriptivist analysis. Secondly, I elaborate on Field (2009) to argue that certain types of disagreement (normative disagreements) involving belief or desire ascriptions reveal an evaluative component. That is, the resistance to dissolve exhibited by this type of disagreement manifests the evaluative nature of the ascriptions; they involve a supportive attitude that cannot be explained if it is assumed that our folk psychological ascriptions are in the business of stating facts. Finally, I argue that central cases of folk psychological ascriptions emerge in contexts where the attributers respond to instances of violation of expectation or counter-normative behaviors with regulative responses, including justifications, exculpations or condemnations. If propositional attitude ascriptions serve to justify or condemn actions, then they are evaluative in nature. Our rationalizations of actions demand to assign commitments and duties to the subject of the action. This practice, I argue, is significantly different from

describing or stating a fact.

These arguments motivate a different approach to the nature of propositional ascriptions. The alternative I canvass in this dissertation is the evaluative view (Chapter 5). The evaluative view propounds that the main function of propositional attitude ascriptions is to assign certain levels of responsibility, merit or demerit to the attributee. Inspired by different anti-descriptivist views (Charlow, 2014; Chrisman, 2012; Hare, 1952; Lance and O'Leary-Hawthorne, 1997), I maintain that propositional attitude ascriptions are not tools for describing or theorizing about others' psychological states. Instead, they burden agents with merit or responsibilities toward a particular content for the purpose of rationalizing, exculpating, justifying or condemning certain actions. Propositional attitude ascriptions help to arbitrate what is permitted or forbidden in our social situations.

A natural consequence of replacing the descriptivist understanding of ascriptions with an evaluative one is that we can specify the social circumstances where mental concepts apply. If mental concepts are discursive devices to burden others with responsibilities, merits and significance, then one must expect to use them in situations which are required to justify or condemn an action, explain counter-normative behaviors, avoiding public sanctions or potential objections, or those which demand regulating or adjusting others' behavior. Thus, social creatures need to engage in those types of situations in order to acquire the conceptual capacity of ascribing propositional attitudes. Children need to engage in conversationally mediated cooperative activities where participants have to adjust and monitor others' participation in order to acquire the capacity to attribute propositional attitudes. In other words, the evaluative conversational hypothesis seems to follow quite naturally from the evaluative approach to propositional attitude ascriptions (Chapter 6).

The evaluative conversational hypothesis is not only well motivated by the alternative to descriptivism I canvass in chapter 5, but also, it is well supported by different sources of empirical evidence in developmental psychology. After characterizing the hypothesis and comparing it with similar views, I present two sources of empirical grounds to support it. Firstly, a solid ground of evidence has shown that conversational interactions with caregivers and siblings strongly correlates with social understanding in children (Jenkins and Astington, 1996; Lewis et al., 1996; Perner et al., 1994; Ruffman et al., 1998). Secondly, many experiments speak in favor of the idea that not all conversational contexts are connected to social understanding. Instead, contexts of dispute, justification and cooperative joint activities are better predictors of folk psychological skills than other conversational circumstances (Foote and Holmes-Lonergan, 2003; Dunn et al., 1991; Dunn, 1994; Hughes and Dunn, 1997; Hughes et al., 2006). The paradigmatic environments to earn those skills are those where others' actions become salient, where our actions are exposed to possible sanctions and where our mutual objectives could raise possible conflictive situations that invite us to recognize others' obligations and duties given the social interaction. Those contexts include playing with siblings, friends and caregivers, solving problems with others, pretending cooperative games, cooperative task resolution and so on. Finally, I face a possible challenge of the hypothesis according to which some empirical findings in animal cognition can serve as counter-examples against the view. If non-linguistic creatures exhibit complex socio-cognitive capacities, for instance, belief attributions, then, one may call the evaluative conversational hypothesis into question. In order to reply this argument, I propose a deflationist interpretation of these findings and I discuss the possibility of justificatory skills in non-human animals.

In conclusion, the evaluative conversational hypothesis depicts an empirical

hypothesis which is well-motivated by several empirical sources in both social and developmental psychology. But also, it naturally follows from a philosophical approach to propositional attitude ascriptions which can avoid several challenges faced by the received view in philosophy of mind.

Resumen

El principal objetivo de esta tesis es ofrecer una hipótesis plausible de la relación entre lenguaje y psicología popular. De acuerdo con esta hipótesis, la cuál denominaré la hipótesis conversacional evaluativa, la comprensión humana de otros agentes en términos de estados mentales requiere ser competente en cierta habilidades lingüísticas complejas. En particular, los seres humanos necesitan participar en actividades conjuntas y cooperativas mediadas conversacionalmente para adquirir la capacidad conceptual de adscribir actitudes proposicionales.

La motivación de una investigación filosófica sobre la relación entre lenguaje y la psicología popular nace del descontento con un supuesto común a todas las teorías empíricas sobre esta relación. En concreto, estas teorías asumen que las adscripciones de actitud proposicional son descripciones de los estados psicológicos de los agentes que se interpretan. Este supuesto da por sentado que los humanos, en tanto que criaturas sociales, necesitan acceder a la maquinaria psicológica de otros agentes por razones de explicación, predicción y coordinación. Nuestras adscripciones de deseos, creencias, esperanzas u otros estados mentales representan o describen estados psicológicos que causan cursos de comportamiento que, de otro modo, nos parecerían extraños. Esta tesis, que introduciré en el capítulo 3 bajo la etiqueta de descriptivismo de la psicología popular, es heredera de una aproximación general al lenguaje de acuerdo con la cual la principal función de nuestras expresiones es describir o denotar diferentes objetos, propiedades y aspectos del mundo (Chrisman, 2007).

Existe una gran diversidad de tesis acerca de la relación entre lenguaje y pensamiento (Capítulo 2). Estas oscilan entre las que niegan cualquier influencia del lenguaje en la cognición social, y otras fuertemente comprometidas con la idea de que el lenguaje es una condición necesaria para las habilidades sociales superiores (Astington and Baird, 2005). A pesar de esta diversidad, todas ellas comparten su compromiso con el descriptivismo de la psicología popular, y de hecho, este compromiso juega un rol fundamental en sus diferentes estrategias argumentativas (Capítulo 3). Mi argumento central es que este descriptivismo de la psicología popular es altamente problemático. Por tanto, el análisis descriptivo de las adscripciones de actitud proposicional debe ser remplazado por un análisis alternativo.

Para dejar claro este punto, presento tres argumentos contra el descriptivismo de la psicología popular (Capítulo 4). El primer argumento presenta diferentes usos de adscripciones de actitud proposicional que son difícilmente interpretables desde un análisis descriptivista. En segundo lugar, elaboro un argumento a partir del trabajo de Field (2009) para defender que cierto tipos de desacuerdos (desacuerdo normativos) que involucran adscripciones de creencias y deseos revelan un componente evaluativo. Esto es, la resistencia a disolverse de este tipo de desacuerdos manifiesta la naturaleza evaluativa de las adscripciones. Dichas adscripciones involucran una actitud de respaldo que no puede ser explicada si se asume que nuestras adscripciones sirven para describir hechos. Finalmente, argumento que algunos casos centrales de adscripción en psicología popular emergen en contextos donde los atribuidores responden a casos de violación de expectativas o comportamientos contranormativos. Estas respuestas regulativas incluyen justificaciones o condenas; y por tanto, si las adscripciones de actitud proposicional sirven para justificar o condenar acciones, entonces tienen una naturaleza evaluativa. Nuestras racionalizaciones de acciones impli-

can asignar compromisos y derechos al sujeto de la acción. Esta práctica es significativamente diferentes de describir hechos.

Estos argumentos motivan una aproximación diferente a la naturaleza de las adscripciones. La alternativa que defiendo en esta tesis es la visión evaluativa (Capítulo 5). Inspirado en diferentes visiones anti-descriptivistas (Charlow, 2014; Chrisman, 2012; Hare, 1952; Lance and O'Leary-Hawthorne, 1997), mantengo que la función principal de las actitudes proposicionales es asignar ciertos niveles de responsabilidad, mérito o demérito al atribuido. Desde este punto de vista, las actitudes proposicionales no son herramientas para describir o teorizar sobre los estados psicológicos del otro. En cambio, las adscripciones asignan a los agentes atribuidos con méritos o responsabilidades hacia un contenido particular con el objetivo de racionalizar, exculpar, justificar o condenar ciertas acciones. Las adscripciones de actitud proposicional ayudan a arbitrar qué acciones están permitidas o prohibidas en nuestras interacciones sociales.

Una consecuencia natural de substituir la comprensión descriptivista de las adscripciones por una evaluativa es que podemos especificar las circunstancias sociales en las que los conceptos mentales se aplican. Si los conceptos mentales son herramientas discursivas que cargan a otros con responsabilidades, méritos y significaciones, entonces uno debe esperar usarlos en situaciones que requieren justificar y condenar una acción, explicar comportamientos contranormativos, evitar sanciones públicas u otras situaciones que demanden regular y ajustar la conducta de los demás. Por tanto, las criaturas sociales necesitan enfrentarse a este tipo de situaciones para adquirir la capacidad conceptual de atribuir actitudes proposicionales. Los niños necesitan involucrarse en actividades cooperativas y conjuntas mediadas por conversaciones donde los paricipantes necesitan monitorizar los comportamientos de los otros para poder atribuir estos estados mentales. En otras palabras, la hipótesis conversacional evaluativa parece

seguirse naturalmente de la aproximación evaluativa de las adscripciones de actitud proposicional (Capítulo 6).

La hipótesis conversacional evaluativa no sólo está bien motivada por la alternativa al descriptivismo defendida en el capítulo 5, sino también está bien apoyada por diferentes fuentes de videncia empírica en psicología del desarrollo. Después de caracterizar la hipótesis, presento dos fuentes de evidencia a su favor. En primer lugar, un lecho sólido de evidencia experimental a mostrado que las interacciones conversacionales con hermanas, hermanos y tutores están correlacionadas con la comprensión social (Jenkins and Astington, 1996; Lewis et al., 1996; Perner et al., 1994; Ruffman et al., 1998). En segundo lugar, muchos experimentos parecen mostrar que no todos los contextos conversacionales está conectados con la comprensión social. Por el contrario, contextos de disputa, justificación y situaciones que involucran actividades cooperativas son mejores indicadores de las habilidades para la psicología popular que otras circunstancias conversacionales (Foote and Holmes-Lonergan, 2003; Dunn et al., 1991; Dunn, 1994; Hughes and Dunn, 1997; Hughes et al., 2006). Los entornos paradigmáticos para adquirír estas habilidades son aquellos donde las acciones de los demás se hacen más relevantes, donde las acciones de uno están expuestas a sanción pública y donde los objetivos mutuos pueden hacer emerger situaciones de conflicto que invitan a reconocer las obligaciones y derechos de otros dada la interacción social. Estos contextos incluyen jugar con otros, resolver problemas con los demás, juegos de simulación cooperativa, tareas de resolución conjunta, etc. Finalmente, encaro un posible desafío para la hipótesis conversacional evaluativa. De acuerdo con este desfío, algunos resultados empíricos en cognición animal podrían servir como contra-ejemplos para mi punto de vista. Si algunas criaturas no lingüísticas exhiben capacidades socio-cognitivas complejas, como por ejemplo, atribuir creencias, entonces la hipótesis conversacional quedaría en

entredicho. Para responder esta posible réplica, propongo una interpretación deflacionistas de estos resultados y argumento contra la posibilidad de que animales no humanos puedan desarrollar habilidades justificativas.

En conclusión, la hipótesis conversacional evaluativa representa una hipótesis empírica bien respaldada por varias fuentes empíricas en psicología social y del desarrollo. Pero además, se sigue naturalmente de una aproximación filosófica de las adscripciones de actitud proposicional que evita varios problemas del punto de vista recibido en filosofía de la mente.

Chapter 1

Introduction

1 A Brief Overview

The aim of this dissertation is to offer an hypothesis of the relation between language and folk psychology[1]: The Evaluative Conversational Hypothesis. According to this view, the capacity for evaluating others in terms of propositional attitudes requires the folk psychologist to develop complex linguistic abilities that allow her to engage in conversational contexts embedded in cooperative projects and joint activities. In these contexts, the courses of action of the different parts become salient and relevant to each other in such a way that the participants need to evaluate and situate each other in a background of permissions and duties. Although this hypothesis is devoted to account for several empirical findings and an important part of my theoretical investigation is to show its empirical plausibility, the motivation behind the view is primarily philosoph-

[1]Several authors prefer to use the concept folk psychology to label the set of all social abilities while reserving the term 'mindreading' or 'theory of mind' to name the specific capacity I am considering here. At the same time, some authors refuse to use the term 'theory of mind' because it is theoretically loaded. In spite of it, I will use the three terms interchangeably here as opposed to 'social cognition' or 'socio-cognitive capacities' that refer to all the social capabilities an agent could have and with 'mindreading' as one of them.

ical. In particular, I contend that the conversational hypothesis is motivated by an evaluative conception of propositional attitude ascriptions. Contrary to the received view about the nature of mental ascriptions, the evaluative view maintains that propositional attitude ascriptions are not descriptions of internal psychological states. On the contrary, propositional attitude ascriptions are evaluations of a person as having different levels of responsibility, merit, demerit or significance towards a particular content. In contexts of explanation, this is translated into situating a person as someone from whom you must expect to behave and speak in ways compatible with the normalized patterns that follow from the content of the propositional attitude. In this sense, mental verbs such as believe, desire or know normally provide information which does not represent any worldly aspect but inform about how an agent guides her behavior, i.e., ascriptions depict different action-guiding information.

The philosophical strategy of exploiting the disanalogy between the different functions of our conceptual apparatus (representing, evaluating, prescribing, or expressing) is not unfamiliar in philosophical inquiry. Different philosophical fields, including the philosophy of language (Prior, 1971; Urmson, 1952), the philosophy of mind (Ryle, 1949, 1979; Wittgenstein, 1953), epistemology (Sellars, 1956) meta-ethics (Gibbard, 2003; Hare, 1952) or meta-epistemology(Chrisman, 2007; Field, 2009) have witnessed some application of this strategy. Despite the popularity, this strategy is barely found in the philosophy of cognitive sciences and psychology[2]. Nonetheless, this alternative has important implications for the debate concerning the influence of linguistic acquisition in the development of folk psychology.

[2]Two exceptions are Boghossian (1990) and Bermudez (2005). However, they consider the non-descriptive approach as an alternative to eliminativism that must be rejected. On the other hand, contemporary philosophers of mind defending similar ideas to the kind I canvass in this dissertation are Gauker (2003) (see also Cleave and Gauker, 2010) and de Bruin and Strijbos (2010)(see also Strijbos and de Bruin, 2012a,b). However, they do not explicitly advocate for the strategy I am embracing here.

In the last decades, a significant dispute concentrating on the nature of the relation between language and cognition has emerged (Carruthers and Boucher, 1998; Gentner and Goldin-Meadow, 2003; Gomila, 2012). This debate has its own counterpart in the development of folk psychological capacities (Astington and Baird, 2005; De Villiers, 2007). Although the dispute seems to be mostly empirical, there is a fundamental philosophical approach behind the different contenders in the debate. I dub this conception folk psychological descriptivism. Folk psychological descriptivism (FP-descriptivism from now on) states that ascriptions of propositional attitudes are states or expressions which function as tools for describing facts concerning the psychological reality of a target. My first aim is to undermine this conception and offer an evaluative alternative.

The evaluative view propounds that mental concepts are discursive devices to burden others with responsibilities, merits and significance in particular situations. This helps us to restrict the domain of application of the concepts to those situations where we need to justify or condemn an action is required, to explain counter-normative behaviors, to avoiding public sanctions or potential objections, or those contexts which demand regulating or adjusting others' behavior. The main consequence of this approach for the relation between language and folk psychology is that social creatures need to engage in these types of situations in order to acquire the conceptual capacity of ascribing propositional attitudes. Children need to engage in cooperative activities conversationally mediated where participants have to adjust and monitor others' participation in order to acquire the capacity to attribute propositional attitudes. When others' behavior become relevant to the achievement of our objectives and cooperative projects, then to burden others with responsibilities, merits and grades of commitment is a useful tool. This hypothesis, I shall argue, has enough empirical motivation and coherence with developmental data to be a serious contender

against other hypotheses about the relation between language and folk psychology.

2 Why Language and Folk Psychology?

Humans spend the majority of their time engaged in social situations carrying out cooperative projects and interacting with each other. This vast amount of interactions would not be possible without a particular kind of skills to deal with social situations. Humans encounter other humans and animals in a way that differs substantially from the way they interact with physical objects –a way that involves an unusual social sensitivity. This unique sensitivity to social interactions is translated into a grade of virtuosity and complexity in cooperation, imitation, cultural learning and other social expertise without comparison in other species. This discontinuity between human and non-human socio-cognitive capacities is usually bound to the human ability to understand agential actions through psychological concepts such as beliefs, desires, fears, hopes and another array of mental terms (Von Eckardt, 1994). This unique human capacity is known as mindreading, mentalizing, folk psychology, or theory of mind. The importance of the inquiry concerning folk psychology is reflected on the increase of empirical research about the topic. As Leudar et al. (2004) say: "the ToM [Theory of Mind] framework has been associated with probably the fastest-growing body of empirical research in psychology over the last 25 years" (pp. 572).

Questions concerning the emergence and functioning of folk psychology are some of the most central in the debates in cognitive sciences and the philosophy of mind in the last decades. The basic focus of discussion concerning folk psychology has been the nature of the mechanisms underlying the process of attribution of mental states to explain and predict behavior. In early debates,

two different views polarized the controversy. On the one hand, the theory-theory view defends the idea that mental state attributions are produced by a kind of theorization based on a systematic corpus of knowledge detailing the connections between perceptual inputs, internal states and behavioral outputs[3] (Baron-Cohen, 1995; Gopnik and Meltzoff, 1997; Gopnik and Wellman, 1994). On the other hand, simulation-theory contends that the process is carried out by different simulation mechanisms based on introspections or off-line sub-personal mechanisms (Goldman, 1989; Gordon, 1995; Heal, 1996, 1998). In later developments, several authors tried to develop certain hybrid versions involving some combination of the processes. In fact, nowadays there is a common consensus about the existence of both types of mechanisms (Carruthers, 2006, 2011; Goldman, 2006; Nichols and Stich, 2003).

As it will be clear during the discussion, this debate is central to any inquiry concerning folk psychological skills. However, this dissertation concentrates on a particular aspect of the acquisition of mindreading, namely, the influence of language on the acquisition of folk psychology. There are different motivations to favor a philosophical exploration of this issue. Firstly, the connection between language and mind has been a focus of philosophical attention that can be traced back to the origins of philosophy. In contemporary philosophy, the relation between language and mind has been a constant the in philosophy of mind and language (Davidson, 1975, 1982; Grice, 1967; Sellars, 1956), but also, in the philosophy of psychology and cognitive sciences (Carruthers and Boucher, 1998; Gomila, 2012). Secondly, given the human proclivity to mentalize and its unique capacity for complex linguistic communication, it makes sense to consider whether there is a conceptual connection between the two capacities. Finally,

[3]At the same time, the theory-theory comes in two versions: the modular view and the theory-formation view. According to the former, implicit theoretical knowledge is integrated into a mental module that is mostly innate (Baron-Cohen, 1995; Scholl and Leslie, 1999). According to the later, the theory is gradually acquired through development (Gopnik and Meltzoff, 1997; Wellman, 1990).

the contemporary debate concerning the relation between language and folk psychology is especially suitable for a conceptual approximation from the philosophy of psychology. On the one hand, in spite of the empirical aspirations of the different parts in the debate, all of them share a fundamental philosophical approach of what ascribing a propositional attitude is (Apperly, 2011, 5) Thus, a genuine philosophical question is which kind of both conceptual and empirical consequences would follow when this conception is replaced. On the other hand, most of the contenders in the debate pursue to be coherent with developmental findings. However, their inclinations to promote a view over others seem to rely on different conceptual views concerning the role of mental concepts, philosophical conceptions of communication, or particular views about the features of language and mind. As a result, there are significant motivations to support a philosophical exploration of the debate between language and folk psychology.

3 Why an Alternative to Descriptivism?

One motivation to seek for an alternative to descriptivism is the discontent with the epistemology behind this view. According to FP-descriptivism, propositional attitude ascriptions are descriptions of internal entities that mirror our internal psychological machinery for the sake of prediction and explanation. Given that folk psychological practice is considered an epistemic practice, we attempt to describe the targets' psychological states because we need to gain knowledge of the causal events that bring out their behavior. In this sense, folk psychological practice is analogous to scientific inquiry. Folk psychologists are somehow as scientists attempting to discover the laws and entities governing behavior (see Knobe, 2010, for a discussion). Prima facie, this picture faces a problem concerning the meaning of mental concepts itself. If verbs such as 'believe', 'know,' 'want' or 'hope' describe internal entities, then only agents with a privileged access to these mental states could fix the true values of expressions

containing these verbs. However, as Ryle (1949) says:

> It was just because we do in fact all know how to make such comments, make them with general correctness and correct them when they turn out to be confused or mistaken, that philosophers found it necessary to construct their theories of the nature and place of minds. Finding mental-conduct concepts being regularly and effectively used, they properly sought to fix their logical geography. But the account officially recommended would entail that there could be no regular or effective use of these mental-conduct concepts in our descriptions of, and prescriptions for, other people's minds. Ryle (1949, 17)

In fact, we know the correctness criteria of these mental concepts, we know how to apply these concepts without any access to others' mental states. Thus, the condition of correctness of them should be different from determining the reference that they are supposed to fix. This motivation is behind the arguments I present in this dissertation, for instance, that FP-descriptivism is incompatible with multiple everyday uses of mental concepts (Chapter 4).

The second motivation starts from an emergent idea among recent contenders of the received view, according to which the primary function of propositional attitude ascriptions is not epistemic but justificatory (Andrews, 2012; Bruner, 1990; Gallagher and Hutto, 2008; Zawidzki, 2013). In other words, mentalizing others is a response to anomalous behaviors that violate the expectations of the attributer. Ascribing beliefs and desires is one of many common responses to counter-normative behaviors framed into the practices of asking for reasons, sanctioning or policing when our actions violate the norms governing our social interactions. The metaphor that emerges from this analysis significantly differs from the analogy of the scientist. Folk psychologists are not like scientists attempting to describe the internal ghostly machinery of others' minds for the sake of explanation and prediction. On the contrary, folk psychologists are as lawyers who advocate for avoiding public sanctions of her clients, or prosecutors who attempt to find the way to condemn them (Zawidzki, forthcoming).

Paradigmatically, ascribing beliefs and desires is practiced in the contexts of excusing or condemning counter-normative behaviors, solving disputes, voicing others, reason explanations and, in general, those contexts where we need to evaluate others for the significance of their actions. In general, ascribing beliefs and desires locates someone in a sphere of responsibilities concerning reasons and motivations for actions. This role of propositional attitude ascriptions in the context of explanation, I shall argue, is incompatible with a descriptivist reading of mentalizing.

4 The Plan

As I said, the main aim of my dissertation is to offer a conversational hypothesis of the relation between language and folk psychology. The primary strategy to motivate this view is to undermine the basic philosophical assumption behind the contenders in the debate, i.e., folk psychological descriptivism. After presenting my evaluative alternative, I shall maintain that the evaluative conversational hypothesis seems to be a plausible consequence of the evaluative framework. To make a case for these ideas, I shall proceed as follows:

In chapter 2, I present the spectrum of hypotheses regarding the relation between language and mindreading. This debate is polarized between two groups of approaches. According to the *communicative views,* no significant role is played by language in mindreading acquisition, or in other words, the only function of language is to communicate. According to the *cognitive views,* different aspects of language are hypothesized as precursors of mindreading acquisition. These precursors include syntax, the semantics of certain concepts, or representational properties. Before presenting those views, I will sketch some empirical findings in developmental psychology that would facilitate the comprehension of the hypotheses and to articulate their main advantages and disadvantages.

The aim of chapter 3 is to present a basic philosophical assumption behind the hypotheses presented in chapter 2: folk psychological descriptivism. According to it, propositional attitude ascriptions describe internal psychological states of the targets for the sake of explanation and prediction. Descriptivism is not only central for the articulation of the hypotheses, but also, it plays a significant role in their argumentative strategies. The main conclusion of the chapter is a conditional claim: if descriptivism turns out to be an inaccurate model of propositional attitude ascriptions, then the arguments supporting the hypotheses presented in chapter 2 are not compelling.

The primary aim of chapter 4 is to undermine folk psychological descriptivism. I shall offer three arguments against it. The first argument seeks to show the incompatibility of FP-descriptivism with many uses of belief and desire ascriptions in everyday contexts. If FP-descriptivism were right, our propositional attitude ascriptions would describe internal entities in the target of the ascriptions. However, propositional attitude ascriptions are not always tied to folk psychological explanation. We use expressions of belief and desire in many other contexts. These uses of mental concepts, I argue, are not always compatible with a descriptivist reading. Furthermore, they seem to have a basic evaluative feature that cannot be captured in the form of descriptive information alone. Thus, we have reasons to motivate the search for an alternative that covers these uses of mental ascriptions as well. Secondly, a consequence of FP-descriptivism is that every situation of disagreement involving ascriptions should dissolve once we clear all the facts up. However, this is not always the case. In several situations, we cannot dissolve our disagreements concerning belief/desire attributions by stating facts. Some of these cases reveal the evaluative nature of our ascriptions; they involve a supportive attitude that cannot be explained if it is assumed that our folk psychological ascriptions are in the

business of stating facts. Finally, I present a recent framework in folk psychology according to which propositional attitude ascriptions are less central to social cognition than the orthodoxy has supposed. Propositional attitude ascriptions are restricted to cases where the attributers respond to cases of violation of expectation or counter-normative behaviors with regulative responses, including justifications, exculpations or condemnations. If propositional attitude ascriptions are used to justify or condemn actions, then they are evaluative in nature. Our rationalizations of actions demand assigning commitments and duties to the subject of the action. This practice, I argue, is significantly different from the practice of describing or stating a fact.

Chapter 5 is devoted to articulate and support an alternative to folk psychological descriptivism: the evaluative view. The rationale behind the evaluative view does not only rely on accounting for the three arguments presented in chapter 4, but also, on different features of the practice of ascribing propositional attitude which reveal its evaluative nature. In particular, ascribing propositional attitudes are usually reactive responses, they have a special connection with action, and they are put into work in explanatory contexts of a special kind. As a result, if the evaluative view is a plausible alternative to folk psychological descriptivism, then we have reasons to assume that the debate concerning the relation between language and folk psychological is somehow defective.

Finally, in chapter 6, I offer an empirical hypothesis about the relation between language and mindreading. If the evaluative view is right, the domains of natural application of mental concepts and ascriptions are those social circumstances where there is a coordinate action conversationally mediated where participants have to monitor and adjust each other's participation. In those contexts, the opportunities for justifying counter-normative behaviors emerge, avoiding public sanctions or potential objections, exercising responsibility and

authority, and shaping conflictive interpretations. Taking this insight on board, a plausible developmental consequence is an evaluative conversational hypothesis: The capacity for evaluating others in terms of propositional attitudes requires the folk psychologist to engage in conversational contexts embedded in cooperative projects and joint activities. In these contexts, the courses of action of the different parts become salient and relevant to each other in such a way that the participants need to evaluate and situate each other in a background of social norms. After articulating the hypothesis, I present a set of developmental findings that speak in favor of its empirical plausibility.

Chapter 2

Language and Folk Psychology

1 Introduction

This chapter aims to offer a theoretical and empirical background to situate the central topics of the dissertation. The primary objective is to sketch a general picture of the different hypotheses of the impact of language on folk psychology. The various views about this influence are divided into two groups. Firstly, the communicative views consider that the only function of language is communication. Thus, there is no significant influence of language (apart from the obvious gaining of information) in the acquisition or functioning of folk psychology. Secondly, the cognitive views claim that language plays a constitutive role in the procurement of folk psychological competence. In section 2, I provide some relevant definitions and the scope of the problem. In section 3, I present some relevant empirical evidence that will help me to situate the empirical plausibility of each hypothesis. In Section 4, I sketch the main hypotheses along with their

major advantages and disadvantages.

Situating the constellation of empirical results and hypotheses helps to pose the contenders in the debate. But also, it helps us to locate the problems of the relation between language and mindreading in a developmental framework that is influenced by different background positions concerning mental architecture or the functioning of mindreading mechanisms. Nonetheless, as it will become clear during the next chapter, these different background positions do not change the common underlying philosophical assumption concerning the understanding of propositional attitude ascription: FP-Descriptivism.

2 Social Cognition and Folk Psychology

As I said in the previous chapter, scientific and philosophical research concerning social cognition has been centered on the ability known as mindreading or folk psychology. Folk psychology is usually understood as the ability to ascribe mental states to others to understand, explain and predict their behavior. For instance, Von Eckardt (1994) introduces the concept as follows[1]:

> Human beings are social creatures. And they are reflective creatures. As such they continually engage in a host of cognitive practices that help them get along in their social world. In particular, they attempt to understand, explain and predict their own and others' psychological states and overt behaviour; and they do so by making use of an array of ordinary psychological notions concerning various internal mental states, both occurrent and dispositional. Let us then consider folk psychology to consist, at a minimum, of (a) a set of attributive, explanatory and predictive practices, and (b) a set of notions or concepts used in those practices (Von Eckardt, 1994).

Those ordinary mental notions are considered the linchpin of human social cognition[2]. Without the capacity to mentalize others, the grade of complexity

[1] Analogous formulations can be found practically in any introductory text to the matter, e.g., see entries by Morton (2009) or Goldman (2012).

[2] In spite of the agreement, an increasing number of dissenters have emerged in the recent years (De Jaegher and Di Paolo, 2007; Gallagher, 2001, 2008; Ratcliffe, 2007; Hutto, 2004; Hutto and Ratcliffe, 2007; Leudar and Costall, 2009). All these scholars share their refusal

and sophistication that our social behavior exhibits would be impossible. For instance, social skills, such as imitation, cooperation or learning seem to be bound to the capacity to attribute mental states (Baron-Cohen, 1999; Mithen, 2000; Dunbar, 2000, 2003).

According to von Eckardt's quote above, folk psychology includes the capacity to attribute a vast range of psychological states, for instance, emotions, perceptions, or intentions. However, many scholars restrict their analysis to the attribution of propositional attitudes. For example, although Nichols and Stich (2003) recognize the role of emotions in our folk psychological capacities, the attribution of propositional attitudes (desires and beliefs) is the primary explanatory target in their view. The centrality of propositional attitudes in the study of folk psychology is reflected in many other authors, for instance, Fodor, who usually refers to folk psychology as "common sense belief/desire psychology"(Fodor, 1987). I shall follow those authors and restrict my investigation to the attribution of desires, beliefs and other propositional attitudes. Propositional attitudes are usually characterized as relations connecting an agent and an object with semantic properties, that is, a representation of the world being in certain ways (Fodor, 1978; Apperly, 2011). Mindreading is generally regarded as the capacity to (meta) represent mental states as such, that is, representing mental states as unobservable representational states causing behavior. In the following chapters, I discuss the assumption behind this conception. However, for the sake of clarification, I keep this definition for now.

As I said in the previous chapter, the different hypotheses concerning the influence of language in the acquisition of those mentalizing skills are polarized between two different views: the communicative view and the cognitive view.

to the centrality that the orthodoxy has assigned to mentalizing in the explanation of social cognition. This set of views is often called interactionism. Interactionists argue that social interactions are usually facilitated by a less intellectualized set of mechanisms than the received views have supposed.

According to the former, language's primary function is to communicate. Thus, language does not influence the acquisition of mindreading faculties. According to the latter, language is a prerequisite for the procurement of the conceptual capacity of attributing mental states to others. Depending on which aspect of language is considered the prerequisite for mindreading, it would make the difference between the different cognitive views.

It helps to clarify the characterization of the different hypotheses if we can present some relevant evidence that elucidates some advantages and disadvantages of the various approaches. Furthermore, those empirical results would help me to test the empirical plausibility of my hypothesis in Chapter 6. Although the main argumentative strategy of this dissertation revolves around the detection of a shared philosophical conception behind the empirical views below, we must not forget that my primary aim is to provide a testable hypothesis about the relation between language and folk psychology.

3 Some Empirical Findings

The empirical evidence I present in this section is not necessarily accepted by all parts in the debate. However, I believe that these studies are sufficiently mentioned and discussed to be considered a good background from which we can appraise the hypotheses. Some of the theses I present in section 4 are not mutually exclusive, and therefore, different compatible views could cover a diverse range of evidence. However, the evidence reviewed in this section could help us to shed light on the explanatory power of each hypothesis. Even if it turns out that we could not isolate a unique linguistic factor that covers the results, it would still make sense to find which hypothesis has more explanatory power.

3.1 The Appearance of Mindreading

A view about the relation between language and mindreading must be coherent with the developmental pattern of emergence concerning the mindreading faculty. For instance, if our empirical record locates mindreading abilities before the appearance of basic linguistic skills, this would automatically rule out any cognitive view regarding the interaction between language and mentalizing. Thus, the hypotheses presented below must account for the empirical results concerning the appearance of mindreading or, at least, provide a plausible explanation of why the hypotheses are not incompatible with these results.

So, when in ontogeny do children acquire mindreading? A widespread consensus in developmental psychology situates the emergence of mindreading capacities around the age of four (Perner and Roessler, 2012; Wellman et al., 2001). Since the beginning, the scientific approaches to mindreading focused on the connection between social cognition and the possibility of ascribing false beliefs (Dennett, 1978b; Harman, 1978). Following this assumption, Wimmer and Perner (1983) designed what now is known as the classical false-belief task (FB-task). In this task, a child is exposed to a scenario where a character, Maxi, puts chocolate into a cupboard x. When Maxi is not present, his mother displaces the chocolate from x into a cupboard y. Children have to indicate the box where Maxi will look for the chocolate when he returns. Only when the child is able to represent Maxi's wrong belief, he is able to point correctly to box x. This task tests whether children have an "explicit representation of the wrongness of this person's belief in relation to one's own knowledge" (p. 103). Wimmer and Perner (1983) found that younger children (3-years old) frequently fail in this task. These results have been consistently reproduced. Furthermore, different versions of the task were proposed (Astington and Jenkins, 1999; Gopnik and Astington, 1988; Perner et al., 1987) For instance, the Appearance-reality tasks

where the child looks at an object that looks like something that it is not, for example, a sponge that looks like a rock. Firstly, the child is asked what the object is and then the experimenter shows what it really is. Then, the child is asked what another friend waiting in another room will think it is.

Consensually, the success in FB-task is the leading indicator of the appearance of full-fledged mindreading in developmental psychology. Wellman et al. (2001) conducted a meta-analysis that seems to demonstrate a robust pattern of coherence concerning the different data involving the false-belief task. They analyzed 178 studies including various factors as age, country of origin, and different versions of the task. The meta-analysis situated a 'conceptual change' that provokes the comprehension of beliefs between the 40th and the 50th month of age. It deserves to mention that children do not merely give the right answer after this range of age, but they systematically give the wrong answer before it. Before passing the FB-task, children are able to ascribe desires by the age of 3. In fact, Bartsch and Wellman (1995) considered three stages in mindreading acquisition: "an early desire psychology, based on non-representational mental state constructs such as simple desires; a transitional desire-belief psychology, in which desires continue to dominate causal-explanatory reasoning despite the existence of an auxiliary concept of belief; and a belief-desire psychology akin to adult understanding" (p. 206).

Before 2005, the emergence of mindreading around four years old was considered the orthodoxy in developmental psychology (Carruthers, 2013; Rakoczy, 2015). Even strong nativist positions (Fodor, 1992; Leslie and Roth, 1993; Scholl and Leslie, 1999) agreed that mindreading was not fully available to agents before this age. Then, Onishi and Baillargeon (2005) published a set of experiments that seemed to show that 15-month-old infants succeed in a non-verbal version of the false-belief task using looking time measure. In the experiment,

the children were exposed to a change-location scenario similar to the Maxi's example presented above. During the experiment, Onishi and Baillargeon measured the looking time of infants to test their reactions. Looking time measure is a standard paradigm in developmental psychology. It works under the assumption that babies will look longer at an event when it violates their expectations. In this case, the infants look longer when the person, who was not present when the object was relocated, picks the object. The infants' sensitivity to false belief was correlated not only with their looking time but also neural responses (Southgate et al., 2007) and helping behavior (Buttelmann et al., 2009; Knudsen and Liszkowski, 2012). Now, this type of experiment is known as implicit false-belief task. Although these findings seem to support the nativist approach to mindreading (Carruthers, 2013; Fodor, 1992; Leslie and Roth, 1993), this interpretation is controversial. Many scholars have provided different deflationist interpretations of these findings(Apperly, 2011; Ruffman and Perner, 2005; Rakoczy, 2015).

A third possibility in the market suggests that FB-task does not show an entire comprehension of others' propositional attitudes. Lalonde and Chandler (2002) argue that children around four cannot understand the subjectivity of the mind, that is, understanding others as representing a particular object or states of affairs from a different aspect or point of view than her own. Rakoczy defines aspectuality as the cognitive aspect we can capture with the distinction between intensionality and extensionality in the philosophy of language:

> Linguistically, the aspectuality of propositional attitudes is reflected in the intensionality of propositional attitude reports. . . .In the context of propositional attitude reports (E.g. "Peter beliefs [sic] that Clark Kent lives next door"), the substitution of co-referential terms ("Superman"/"Clark Kent") is not truth-value-preserving ("Peter beliefs [sic] that Superman lives next door" can be false even though "Peter beliefs that Clark Kent lives next door" is true). Crucially, aspectuality is not just an accidental or peripheral but an absolutely fundamental and essential property of beliefs and other propositional attitudes: there is no grasp of what propositional attitudes are with-

> out some basic grasp of their aspectuality (Rakoczy, 2015, 4).

Propositional attitude ascriptions introduce intensional contexts, where two co-referential terms cannot be substituted *salva veritate*. Now, Lalonde and Chandler (2002) defend a later age for the acquisition of mindreading (around 7). The empirical support for this proposal relies on empirical evidence concerning the detection of aspectuality or intensionality (Apperly and Robinson, 1998, 2003; Kamawar and Olson, 2009, 2011; Russell, 1987; Sprung et al., 2007). The structure of these experiments was to present children with a story or premise where an object has two aspects (A and B). A person does not know that they are the same object and she is looking at it under the aspect A. Then, the children were asked whether or not the person is looking at B (The correct answer was no). Children between 6 and 8 years old find the task difficult to understand. These findings are used to locate the understanding of others' mind later than it was supposed.

In any case, it is undeniable that any theory concerning folk psychology acquisition should be coherent with one of these sets of results situating the appearance of mindreading around the 13 months, 4 years or 6-8 years. If language plays any role or not in the acquisition of mentalizing, this should impact in the developmental path of full-fledged mindreading. Whatever your position concerning language and mindreading is, it must be coherent with one of these patterns of empirical data, and also, it should provide reasons to discard or reinterpret the other results.

3.2 Vocabulary, Caregivers and Mindreading

There are different types of experiments in developmental psychology exploring the connection between language and folk psychology. One way to approach this connection is what Carpendale and Lewis (2006) coined as 'language as

a window to development'. This term labels a set of experiments that take children's linguistic uses of mental vocabulary as a reliable marker of the way they construct their understanding of the mind. However, it is more interesting to concentrate on correlation and training studies that point out to causal connections between specific linguistic abilities and mindreading.

Our first focus of interest is the wave of experiments showing the correlation between exposition to vocabulary and linguistic contexts, and the mastery of false-belief tasks (Astington and Jenkins, 1999; Cutting and Dunn, 1999; Happé, 1995; Jenkins and Astington, 1996; Ruffman et al., 2002; Watson et al., 2001). These studies point out that parents who reported discussing and elaborating on mental states about different situations to their children tended to have children who were advanced in passing the false-belief tasks. For instance, in longitudinal studies, Dunn et al. (1991) analyzed the interactions of 50 children with their families when they were 33 months. After that, the children were tested about their belief understanding when they were 40 months old. The findings suggested a correlation between the exposition to mental vocabulary and understanding beliefs when tested. Other studies show the influence of the use of mothers' mental vocabulary and the folk psychological comprehension of children (Dunn and Munn, 1987; Furrow et al., 1992; Moore et al., 1994). For instance, Ruffman et al. (2002) designed an experiment to evaluate children's linguistic competence and social understanding, and mothers' tendency to talk about mental states. They found that mothers' use of mental vocabulary predicted children's understanding of mental states and linguistic abilities. The main problem of these findings is that they demonstrate correlation but not causal connection. Given this, one may interpret that a mother's uses of mental vocabulary can be due to the attunement of the mother when conversing with her child. So, the increasing of mental vocabulary could be produced because

the child is already more skillful in social understanding and that is reflected in her mother's language. However, Astington and Jenkins (1999) ruled out this possibility. They found that success in the theory of mind was not a predictor of linguistic abilities. Rather, the reverse was true: language proficiency was a good indicator of improvement in the theory of mind. Finally, Happé (1995) investigates the correlation between lexical knowledge, measured by standardized tests as the Peabody Picture Vocabulary, and success in false belief tasks. The studies included preschoolers, autistic children, and mentally handicapped subjects, but the correlation was found only in the two former.

What these empirical results show is a strong correlation between both the exposition to a certain vocabulary and general lexical knowledge and children's understanding of mindreading abilities. The results do not clarify the nature of this relation; in principle, it is not clear whether or not the influence of the mother's linguistic skills is due to the use of specific vocabulary or the nature of the engagement itself. Furthermore, the correlation between children's linguistic competence and FB-task demonstrates that it is the children's capacity itself what correlates with mindreading skills. Thus, the influence of the exposition could be interpreted as a way of facilitating the comprehension of linguistic skills that, in turn, pave the way to mindreading acquisition. In any case, these experiments impose another source of evidence which will help to clarify the views I present in the following sections: any view of language/mindreading relation must explain why being exposed to certain conversational exchanges with mental vocabulary and acquiring general linguistic knowledge correlates with FB-task success.

3.3 Perspective-shifting and Complementation

Another group of correlational studies attempting to demonstrate the relation between language acquisition and success in the false-belief task are the experiments performed by Jill de Villiers and colleagues (De Villiers and De Villiers, 2000; de Villiers, 2005; De Villiers and Pyers, 2002). They show a strong correlation between the mastery of sentential complement and mindreading. In the experiments, the researchers measure the success of the subjects on linguistic tasks including sentential complements. For instance, the scholars confront the subjects with questions such as the following:

> *The Mom said she bought apples, but look, she really bought oranges.*
> *What did the Mom say she bought?*

De Villiers and colleagues demonstrate that children who answer this type of question correctly succeed in the false-belief task more than children who do not.

These results were complemented with training studies. For instance, Hale and Tager-Flusberg (2003) found that training children on sentential complement exercises improves their scores in false belief tasks. In these experiments, 60 children who failed to perform the FB-task and complementation test were trained in different tasks including false belief, sentential complements, and relative clauses (control). After that, they were tested again in FB-task, complementation task, and relative clauses. Children trained in sentential complements improve significantly in the FB-Task. In contrast to the other results, those data strengthen the hypothesis of a causal connection between syntax and mindreading. Furthermore, they found that complementation with communicative verbs (*tell, say*) produced the same effects as mental verbs (*think, believes*).

A fair question concerning these studies is if they indicate a special connection of mindreading with syntax or they just reflect a symptom of a more general

linguistic capacity. Lohmann and Tomasello (2003) designed other training task experiments issuing related results for this question. The experiment consisted of four different training conditions involving a deceptive object (A pen that looks like a flower). The first training comprises talking about the deceptive nature of the object using communicative or mental verbs (full training). The second, talking about the deceptive nature of the object without communicative or mental verbs (discourse-only training). The third, showing the nature of the deceptive object without linguistic help (no-language training). The fourth, talking about the object without highlighting its deceptive nature using communicative and mental vocabulary (sentential-complement-only training). After the training, the subjects of the experiments improve significantly in the false-belief tasks after they perform the full training, but also in discourse-only training and sentential-complement-only training. These findings indicate complementation matters but also the perspective-shifting discourse. Taking both sets of results together, it seems fair to conclude that complementation studies reflect a general important aspect of language for mindreading, namely, aspectuality. The capacity of using embedding sentences may influence mindreading acquisition as part of a general comprehension that we can represent a state of affairs from different perspectives (see also Gomila, 2012, 87).

3.4 Conclusions

These empirical findings in developmental psychology reveal a complex relation between language and folk psychology. Most of the experimental results seem to indicate that sophisticated linguistic communication is somehow related to the capacity for attributing propositional attitudes. However, the different experiments point out to different ingredients of language that could be important or necessary for a normal development of human socio-cognitive capacities. In any

case, the studies establish different ways of testing the empirical plausibility of the theories about the influence of language on folk psychology. Even if a theory neglects the influence of language, then it must provide an alternative story of what is going on in these studies. The next section aims to explore the different theories about language/mindreading relation and their main advantages and disadvantages.

4 Language and Mindreading: Five Hypotheses

The different experimental results reviewed above demonstrate the complex interrelation between language and folk psychological capacities. Given the difficulty to pin down the nature of the relation, it is not surprising we can find a great variety of theoretical products to buy in the market. This section aims to review this variety of contenders. For the sake of clarity, I continue with the division between the Communicative views and the Cognitive views. The literature concerning the impact of language on cognition broadly considered has increased significantly (Gentner and Goldin-Meadow, 2003; Gomila, 2012). However, through this chapter, I will restrict the discussion to those theories and approaches that focus on the impact of language on social cognition. The distinction between communicative and cognitive views is quite straitjacketed. In principle, there is a continuum of possibilities where it is expectable to find different views and approaches that could be difficult to recognize under one of these labels. Furthermore, several of those views are not mutually incompatible, so the influence of language and mindreading can be found at different levels. A more vivid picture of the map of views I will discuss is a set of different focal ideas that we should pay attention to in order to understand the scientific and philosophical approximation to the connection language/mindreading. Accordingly, the location of the authors discussed below under one of those terms must

be considered as a dialectical strategy for expositive purposes.

4.1 The Communicative Views

In general, the communicative views consider language as a peripheral capacity of the mind, that is, language is separable from the rest of the cognitive endowment. Language is a cognitive mechanism whose function is expressing information brought from the mind and processing information received from the outside. Apart from this, there is no cognitive function that language may play in the development of mindreading capacities. So, the communicative views contend that language plays no significant role in the acquisition or performance of attributing propositional attitudes. Of course, what one may consider significant or not is a question of degree. Thus, I divide the views into two groups. Firstly, the strong version of the communicative view claims the mindreading mechanisms are entirely independent of our language faculty. Secondly, the weak versions of the communicative views maintain that language may play some role in the acquisition of mindreading capacities, but this influence is basically negligible. To see the contrast with a genuine cognitive view, the weak version would deny that language is a pre-requisite to acquire mindreading in any sense or play any constitutive role in the 'conceptual change' that acquiring mindreading supposes. However, it accepts that it could improve our access to certain information required for the development of mindreading abilities.

The Strong Version

As I said before, the communicative views claim that language is not required to develop the theory of mind. There are two primary reasons to support a strong communicative view. Firstly, one may argue that language cannot influence mindreading by expanding upon the idea that mindreading capacities are

basically innate (Baillargeon et al., 2010; Carruthers, 2013; Fodor, 1992; Leslie, 1994). A way to support the view is considering the results produced by the experiments concerning implicit versions of the FB-tasks reviewed above. Given the precocity of the subjects able to pass the implicit FB-task, it is fair to claim a nativist approach to mindreading capacities. Now, opting for a strong version of the communicative view carries the load of giving a plausible explanation of why 4 years old children, who are supposed to be competent mindreaders, fail to pass explicit versions of the tasks. The primary answer exploits the idea that explicit FB-tasks results reflect different problems regarding the verbal conditions of the tasks. For instance, some authors have argued that FB-tasks demand more working memory space that children can afford (Fodor, 1992; Leslie and Roth, 1993; Scholl and Leslie, 1999). Evidence supporting this idea correlates success in the task with factors such as inhibitory control (Carlson et al., 1998; Leslie and Polizzi, 1998), memory (Freeman and Lacohée, 1995), or executive functions (Carlson and Moses, 2001). Other results backing the nativist view come from studies showing that children around 3 years old can pass simplified versions of the explicit FB-task (see for instance, Rubio-Fernández and Geurts 2013).

A second reason to support the strong communicative version may rely on a particular picture of linguistic communication which starts from the idea that communication requires certain mindreading skills. This strategy exploits a widely spread conception of communication according to which linguistic communication is not only a question of de-codifying information linguistically encoded, but also, a question of inferring the intended meaning of the speaker. In those so-called neo-Gricean approaches to communication, the inferential capacities that capture the intended meaning necessitate mindreading abilities. The process of understanding meaning starts with the de-codification of the content

coded in an utterance (*linguistic meaning*). However, this codified meaning is only an input in the process. In order to understand the entire speech acts (*speaker's meaning*), the hearer has to infer contextual information. Now, the question is how this information is inferred. According to this picture, the utterance does not only convey conceptual information but also it triggers certain expectations that help the hearer to track the speaker's meaning. These expectations are triggered when the listener can understand that the speaker intends to communicate something with her utterance and that he has the intention to inform the hearer that he has the intention to communicate (Carston, 2002; Sperber and Wilson, 1996)[3]. These expectations plus the codified information help the hearer to infer the most relevant contextual information in order to grasp the speaker's meaning. Mindreading capacities are not only necessary in the process of interpretation to attribute intentions, but also, they are required in the process of speech production in order to choose the appropriate words to provoke the intended reaction in the hearer.

Some authors have exploited this conception of communication in order to argue for a strong communicative version (Fodor, 1998; Leslie, 1987). Rather than playing a role in the theory of mind development, complex communication through language cannot be performed without the functioning of mindreading mechanisms. Complex linguistic communication relies on intention understanding and, therefore, in the capacity of understanding each other in mental terms. Thus, communication is only possible when mindreading is acquired instead of the other way around (for a different solution see Fernández-Castro, 2015a). Notice that this picture of communication would reinforce the nativist position

[3]Wilson and Sperber (see 2012) have recently modified his view concerning this point in order to accommodate some of the evidence regarding the appearance of full-fledged theory of mind. According to their revised version, the model only required the capacity to attribute attentional states or a pragmatic module that drive the process. From the modified version, Mindreading capacities would be only necessary for complex linguistic phenomena such as irony or sarcasm.

concerning what is going on in the explicit FB-tasks. It is expectable children's working memory or executive functions may fail at passing the task when they have to keep track of the false belief of the Muppet while reading the experimenter's intentions when formulating the questions. However, it is hard to see why training in complementation, shifting-discourse or being exposed to a certain type of vocabulary would improve working memory and executive abilities more than any other linguistic or non-linguistic capacity that exercise those functions.

The Weak Version

Another possibility in the communicative side is to recognize that language could be a useful tool for acquiring folk psychological capacities but it is not a prerequisite. On this account, language is understood as a vehicle which carries information that could improve the access to the knowledge required to gain mindreading abilities. However, the lack of linguistic abilities does not necessarily involve interrupting the developmental acquisition of mindreading. This perspective seems to be mostly advocated by some versions of the theory-theory (Bartsch, 2002; Gopnik and Meltzoff, 1997; Gopnik and Wellman, 1994; Wellman, 2014). These versions of the theory-theory claim that the developmental patterns of mindreading acquisition exhibit a pattern of acquisition similar to the development of scientific theories. According to this view, children construct theories of the world, and they alter and revise such theories based on new evidence. Children accumulate social information through being exposed to the social environment where interactions among different agents reveal the way to think about them as minded. The developmental changes in children's expertise in theory of mind reflect change in the quasi-paradigm they use to understand others:

> In our proposal, the change to understanding representational states of mind is viewed as a development within children's theory of mind. Conceiving of the distinctive representational nature of some mental

> states comes about for children via their struggles to understand mental states at all (e.g., desires, emotions, and so on) and their emerging conception of an internal world of mental contents separate from the real world of occurrences or states of affairs (Bartsch and Wellman, 1995, 194).

Drawing a parallel with scientific understanding, Bartsch and Wellman consider that children only incorporate a proper representational understanding of belief at the age of four, when they use belief attribution to provide a causal-reasoning explanation. Children's theory of mind is refined with the evidence children confront during development:

> The [scientific] theory-theory proposes that there are powerful cognitive processes that revise existing theories in response to evidence. If cognitive agents began with the same initial theory, tried to solve the same problems, and were presented with similar patterns of evidence over the same period of time, they should precisely converge on the same theories at about the same time. These assumptions are very likely to be true for children developing ordinary knowledge (Gopnik 2003, 248, see also Gopnik and Meltzoff 1997, 52-53).

So, children need to be exposed to social information in order to acquire a sophisticated theory of mind. It is in social interactions where they received the necessary information to obtain the implicit theory that they use for prediction and explanation. The non-nativist view could accept the findings presented above without recognizing a constitutive connection between language and folk psychology. For instance, arguing that language is only a mere instrument to convey social information required for children to revise their proto-scientific theory.

Perner and his colleagues defend a similar position concerning the different competences that children's develop for reasoning about the mind. However, rather than an advance in the sophistication of the theory about the mind, Perner claims the developmental path exhibits a conceptual change in the conception of behavior. Before four years old, we must understand the "child as a situation theorist who is not yet capable of metarepresentation. That is, young

children can represent different situations, real and imagined, but have no conception of something representing these situations" (Perner, 1991, 215). After four years old, they develop metarepresentational abilities; they recognize the existence of basic representational mental states they use to acquire a comprehensive causal explanatory theory of mind. The core developmental step is that understanding beliefs "depends on the ability to realize that things in the world (states of affairs, external referents, or –we now like to use the most neutral terminology—"targets") can be represented as being different than they are (known to the child)" (Perner et al., 2005, 221) In spite of the differences, this perspective shares with the theory-theory its lack of commitment to a strong influence of language in the theory of mind. As Perner and colleagues say: "we have no great theoretical commitment on this point [causal role of language] and see the link by default in the role for language as provider of information that is required for building a theory of mind" (Perner et al., 2005, 222)[4].

According to the weak view, children experience a conceptual change around four, which explains the success in explicit FB-task. Furthermore, Perner and his colleagues (Perner and Roessler, 2012; Ruffman and Perner, 2005) have provided a plausible interpretation of the implicit FB-task success. According to them, this success does not require a full mastery of propositional attitude ascriptions. Instead, children can pass these tasks with a 'situation theory' or 'teleological explanation': "Why does the baker get up at 3 a.m.? Well, the bread needs to be ready by 6 to go to the super-markets, and it takes that long to bake. This is a humble example of a teleological explanation: it makes the

[4]At this point, one may have noticed that both versions of the communicative views seem to be committed to a version of the theory-theory. While the strong communicative views' defenders are usually modularist theory-theorist, the weak views' defenders are non-nativists. This connection is purely accidental. For instance, a nativist could argue for a weak view by considering that exposition to natural language may trigger the development of a theory-theory module whose developmental pattern is otherwise pre-programmed. Furthermore, one may argue for a communicative view from the simulation-theory side. For instance, considering that our simulation mechanisms are developmentally isolated from natural language.

baker's unusual behavior intelligible not by appeal to his mental states such as his desire to make bread etc., but in terms of the objective reason-giving facts of his situation" (Roessler and Perner, 2013, 35); I will discuss this type of explanations in terms of reason in chapters 5). Children before 4 can pass implicit FB-task because they can anticipate the target on the basis of objective reasons concerning the situation, rather than on the basis of mental state attribution. Regarding vocabulary and general linguistic competence, the weak communicative view can defend that language facilitates the transition by providing more information about others' mind. This would explain the experimental evidence concerning the exposition to mother-children interactions. Similarly, this view could account for the training task involving discourse about appearance/reality since training in these tasks would provide more information about the subjective perspective of the others.

4.2 The Cognitive Views

In the previous sections, I defined the cognitive views, roughly, as the positions according to which language plays some role in the functioning or acquisition of theory of mind. These positions can vary according to different grades, parameters and commitments. In order to make sense of the constellation of theories under the label 'cognitive view', I divide the theories depending on the feature of language they take as playing the central role in the influence to acquire mindreading abilities. Before discussing the different perspectives, it is worth highlighting that these positions are not mutually exclusive. One may defend that language influences theory of mind at different levels.

The Syntactic View

In the previous section, I mentioned the empirical evidence provided by de Villiers and colleagues on the correlation between the training in sentential

complements and success in false-belief task. On the basis of those results, these authors hypothesize that this syntactic ability is a necessary condition for the acquisition of a full-fledged theory of mind:

> In essence, then, the claim is that once the child has the grammatical machinery in place to represent a false complement, then this opens up the possibility of false belief reasoning. Before the possession of the appropriate grammatical machinery and key vocabulary (such as mental state verbs, 'believe', 'think', etc.), children may have a range of important understanding of both their own and other people's mental states, but the explicit understanding of the content of false beliefs is not possible (de Villiers and de Villiers, 2003, 170).

There seems to be a causal relation between complementation and the acquisition of theory of mind. According to de Villiers and de Villiers (2000), the complementation hypothesis explains training correlation but also the appearance of mastery in false belief task around 4, when children start to use complementation properly. Furthermore, the hypothesis coheres well with another fact concerning mindreading: the appearance of desire talk before belief talk. Bartsch and Wellman (1995) demonstrated that children's desire talk, including desire expressions and explanation regarding preferences, appears before they express beliefs. De Villiers and De Villiers (2000) argue that this happens because desire talk is not always tied to complementation. For instance, we use expressions such as 'I want an apple'.

Notwithstanding the results, the syntactic proposal encounters a significant problem. According to the hypothesis, children immersed in natural languages where desire talk can be used without complementation (Spanish, English) should acquire vocabulary about desires before vocabulary about beliefs. On the other hand, one may expect the vocabulary about desires and beliefs to appear at the same time in subjects immersed either in languages where both vocabularies require complementation or where neither of them requires complementation. The problem is that both cases have been proven to be false. Firstly, Tardif and Wellman (2000) tested whether or not Mandarin and Cantonese-

speaking children use desire talk before belief talk. Mandarin and Cantonese have a simple grammatical construction to talk about desires and beliefs, so one must expect these children to talk about desires and beliefs at the same time. However, as in the English speakers case, Cantonese and Mandarin-speaking children talk about desires before they talk about beliefs. Perner et al. (2003) tested German and Austrian-speaking children (German and Austrian requires complementation for both desires and beliefs) with similar results.

De Villiers (2005) reacted to these problems by reformulating their hypothesis. She introduced the distinction between *realis* and *irrealis* expressions. Following Bickerton (1981/2015) realis verbs are those which are complemented with observable objects or state of affairs, while irrealis verbs are complemented with unobservable or no-real objects or states of affairs. For instance, examples of realis constructions are expressions such as 'I will kick the ball' or 'I bought ice-cream', while examples of irrealis constructions are 'I want to go' or 'I should finish my homework'. According to de Villiers, while children can perceive 'want' as irrealis, they take 'think' to be realis. This distinction may help to avoid the problem of the desire talk. The point is that while desires are conceptually connected to non-observable or no real objects because of the direction of fit –that is, because they imply to refer to things that still are not the case– beliefs plus complementation require exposition to notice the possibility that realis verbs can be attached to false clauses. This difficulty would explain the lag between talking about desires and beliefs. The basic idea behind the distinction, de Villiers argues, is that realis verbs plus complementation (Think) mark 'points of view': "we can consider subject PoV [point of view] to be the distinctive feature dictated by this subclass of say/think realis verbs" (de Villiers, 2005, 211).

I think there are at least two objections to this solution. One is that the

realis/irrealis distinction is not sharp. If the verb 'want' is supposed to be irrealis, then it could not mean that the object is not observable or actual, because one may desire something that one can observe and that is real. Furthermore, it could be argued that the understanding of desire involves understanding the point of view of the attributee as well. For instance, someone's preference can be radically different from my own. On the other hand, one may object that it is not so clear in which sense we can say that the verb 'think' is realis, given that the object of the verb can be false. The second objection is that this version of the hypothesis betrays the spirit of the syntactic hypothesis because the realis/irrealis distinction does not seem a syntactic distinction. Additionally, if the point of the distinction is to show different points of view, it could be possibly argued that other features of language can help to appreciate them. In fact, these changes in the position seems to indicate that syntax is insufficient to account for the impact of language in mindreading acquisition. In particular, the distinction between realis/irrealis points out to a semantic ingredient as a way to complement the theory (Gomila, 2012, 87). This seems to point out in the direction I discussed above, complementation, as perspective-shifting, indicates a general understanding of the capacity of representing certain events from different aspects or points of view.

The Semantic View

A second cognitive view considers that the acquisition of mental terms determines folk psychology acquisition (Olson, 1988; Segal, 1998; Smith, 1996). According to this view, our folk psychological capacities are bound to the linguistic ability of using verbs such as want, believe or think in conversations. For instance, Olson (1988) claims that children exposed to this mental vocabulary "acquire the cognitive machinery that makes intentional state ascription literally true of them . . . Thus the behaviorists may be correct in denying the

reality of beliefs and desires to infants, while the intentionalists may be correct in claiming the reality of beliefs and desires in older, linguistic children and adults" (p. 420). In other words, acquiring mental concepts implies to acquire the capacity to understand others as possessing mental states bringing out their behaviors. The main difference between the semantic view and the weak version of the communicative view above is that language contribution is not a question of acquiring information that helps us to determinate that others have beliefs or desires that cause their behavior. On the contrary, the semantic view claims that linguistic understanding of the terms such as beliefs and desires provides the accurate conceptual structures that allow us to engage in mental state ascriptions. According to Segal (1998), representing someone's beliefs or desires requires to (meta)represent others' attitudes to a representational content. These meta-representations need a particular kind of conceptual abilities, namely, the capacity of representing relations between agents and propositions. This particular representational capacity cannot be acquired without acquiring the concepts of beliefs and desires which he considers to be language-dependent. Language is not only a vehicle for information, but it provides the required structure to represent others' attitudes towards a mental content. The semantic view is well situated to explain the correlation between the mental talk of caregivers and the success in FB-task of children. By exposition to these terms in conversational contexts, children acquire the appropriate conceptual structure to understand others as possessors of mental states that mediate between their perceptual inputs and behavioral outputs. Furthermore, the semantic view coheres with the orthodoxy concerning the appearance of full-fledged mindreading around four years old, just after children start to use mental verbs properly (Bartsch and Wellman, 1995). However, the studies of Lohmann and Tomasello (2003) present a challenge to the semantic view. On

the one hand, the discourse-only training improves the score in the FB-tasks without using mental predicates vocabulary. On the other hand, exposition to mental predicates does not make a difference comparing with communicative verbs such as 'say' or 'tell' (complementation training). As I mentioned before, training in complementation with communicative verbs improves the success in false-belief tasks as much as training children explicitly in the task or in complementation with mental verbs (see also, Hale and Tager-Flusberg, 2003).

Second-order Dynamics View

The semantic view defends that the kind of meta-representational capacities that an agent necessitates to reason about others' mental states are gained via mental state verbs. Another way to approach the acquisition of these meta-representational capacities is through the general capacity of language as representational vehicle. Several philosophers (Clark, 1998; Bermudez, 2003a,b; Jackendoff, 1996) have argued that language facilitates human cognition because it enhances the capacity of having access to our own thoughts (for a criticism of this approach see Fernández-Castro, 2017). Language exhibits features that make it an appropriate vehicle to codify thought so that it can be object of farther thoughts. In other words, language is what Clark (1998) calls a facilitator for second-order dynamics. Language has an appropriate representational character to be recruited for cognitive purposes. For instance, language essentially codifies thoughts that can be made available for consciousness. Although these philosophers share this view of language as cognitive facilitator, they differ about what properties are relevant to make language an appropriate vehicle for conscious access. Clark claims that language is the appropriate vehicle for second-order dynamics because of its context-independency and neutral-modality:

> [Language is] a type of code which minimizes contextuality (most words retain more-or-less the same meaning in different sentences in which they occur), is effectively modality-neutral (an idea may be prompted by visual, auditory or tactile input and yet be preserved

> using the same verbal formula), and allows easy rote memorization of simple strings (Clark, 1998, 178).

According to him, regular cognition is amorphous and context-dependent. Anchoring sentences in our working memory makes thoughts context-independent and introspectable. That is what makes language the social scaffolding for cognition. On the other hand, although Jackendoff and Bermudez share a similar view about the general function of language, they focus on different linguistic features that underlie this function. According to Jackendoff, linguistic items are adequate for conscious access because of their phonetic nature, since all consciousness is perceptual[5]. For our purpose, the most interesting version of this position is held by (Bermudez, 2003a,b) Bermudez explicitly maintains that second-order dynamics is what facilitates the acquisition of certain mentalizing capacities, namely, the attribution of propositional attitudes:

> Folk psychological reasoning is a paradigm of meta-representational thinking, where metarepresentational thinking involves thinking about thoughts—taking thoughts as the objects of thought, attributing them to other subjects, evaluating their inferential connections with other thoughts, and so on... One might argue, for example, that thoughts must be vehicled in a way that is consciously and reflectively accessible if they are to feature in metarepresentational thinking, and that the only possible vehicles are linguistic. (Bermudez, 2003b, 35).

Bermudez argues that language is the only representational vehicle that allows us to access propositional thought. He posed two reasons why natural language is the best option to be considered the appropriate vehicle for metarepresentations. On the one hand, the kind of metarepresentational reasoning involved in folk psychology requires having an appropriate structure for reasoning. However, cognitive maps or images, the usual suspects to carry the representational content in non-linguistic animals, do not exhibit the necessary structure for in-

[5] According to Jackendoff (1996), humans have conscious access to intermediary level of representation which includes phonetic representations. This is the reason why inner speech representations are necessary for consciously accessing our own thoughts. These and other functions of inner speech have been extendedly discussed (see Martínez Manrique and Vicente, 2010, 2015, for review and discussion. For alternative approaches to inner speech see Jorba and Vicente, 2014; Fernández-Castro, 2017)

ferential reasoning. On the other hand, the required vehicle needs to be at the personal level, that is, to be consciously accessed. This eliminates the other candidate with the appropriate structure: the language of thought (Fodor, 1975).In conclusion, according to the second-order dynamics view, language is a prerequisite for mindreading because it is the appropriate representational vehicle to engage in the meta-representational skills required to understand others in terms of beliefs and desires.

Notice that the primary motivation behind the second-order dynamics view is conceptual. However, the second-order dynamics view counts with some empirical plausibility. From this view, one may expect children who engage more in conversational contexts to be more competent in socio-cognitive understanding because they will be more competent in engaging in linguistically coded thoughts. This coheres with some of the lexical findings correlating general language exposition and competence. Furthermore, this also fits quite well with the orthodoxy in developmental psychology situating the acquisition of mindreading after 4, when children exhibit a sophisticated linguistic competence.

The problem of the approach comes with the other sources of evidence. According to the second-order dynamics view, the impact of language on cognition is related to the general role that language plays for conscious access to thought. However, conscious access does not seem to be necessary for displaying mindreading abilities or, at least, it does not seem to play an important function for its acquisition[6]. In principle, the reviewed empirical evidence points out to other types of connections. For instance, the necessity of language for acquiring the appropriate cognitive structures that permit attributing mental states. Furthermore, there are two pieces of evidence that the view seems to have problems to accommodate. On the one hand, if reasoning about mental

[6]In fact, the mindreading faculty could play a similar role (Martínez Manrique and Vicente, 2008)

states only requires general linguistic abilities in order to consciously attend to linguistic structures, then one should not expect specific syntactic structures to correlate with FB-task success. On the contrary, one should expect a progressive increasing of socio-cognitive capacities along the increasing of general linguistic capacities. Similarly, one may expect any type of discourse to correlate with mindreading capacity, and not particular type of conversations as those involving perspective-shifting or exposition to a certain type of vocabulary.

4.3 Recapitulating

The aim of section 4 was to review several approaches to the relation between language and folk psychology. Those approaches present both advantages and disadvantages to account for some phenomena concerning the acquisition of mindreading skills. For instance, while the cognitive views seem to be better positioned to account for the explicit FB-task, the communicative views have a better approach to the implicit FB-tasks and some interesting ways to account for the failure of children in the explicit one. But while testing the empirical plausibility of each account gives us some idea of possible arguments for or against any of these theories, this is not the main strategy I will follow in this dissertation. Nevertheless, it gives us a first glance of what is going on in the philosophical and scientific arena of the relation between language and social cognition.

5 Conclusions

There is an interesting set of theoretical positions in the market for attempting to elucidate the contribution of language to the acquisition of mentalizing abilities in humans. Presenting some of the main advantages and disadvantages of those approximations regarding some scientific results was the aim of this chapter.

However, this will not be the primary motivation to discard those views and promote my own. On the contrary, I will do it on the basis of the philosophical assumption concerning the nature of propositional attitude ascriptions. In the following chapter, I present such philosophical assumption and the role it plays in the articulation of the hypothesis I have presented here.

Chapter 3

Folk Psychological Descriptivism

1 Introduction

The aim of this dissertation is to explore the consequences of a particular treatment of propositional attitude ascriptions for the debate concerning the role of language in folk psychology. The reason to motivate such treatment is to seek an alternative to the received philosophical analysis of propositional attitude ascriptions that, as I will show, is shared by the different positions reviewed in the previous chapter: folk psychological descriptivism. Folk psychological descriptivism is the thesis according to which propositional attitude ascriptions are state-descriptions expressing genuine predicative judgments that denote or stand for states of the world, namely, internal states mediating between perceptual inputs and behavioral outputs. In this chapter, I characterize this assumption, and I attempt to demonstrate how the different conceptions of folk psychology have embraced it. After this, I defend that descriptivism plays a central role in

the articulation of the various arguments supporting the hypotheses concerning the language-mindreading relation. The conclusion of this chapter is that if descriptivism turns to be wrong (as I will argue in chapter 4), then the arguments supporting the different hypotheses would become deflated. Descriptivism is a philosophical approach. However, it influences scientific research concerning mindreading including the various positions in developmental psychology and the relationship between social cognition and language.

2 Folk Psychological Descriptivism

A visible idea in the different disputes concerning the nature of folk psychology is that the output of folk psychological capacities are descriptions of psychological states (Bartsch and Wellman, 1995, 4-5; Botterill, 1996, 115; Fodor, 1998; Goldman 2006, 100; Gopnik, 1996, 187; Gopnik and Meltzoff, 1997, 13-42; Leslie, 2000, 207-208; Leslie and Thaiss, 1992, 231; Perner, 1991, 38-40; Wellman 1990, 9-10). This idea is widely accepted in the debate concerning the nature and development of folk psychological mechanisms. All parts in the discussion seem to share this basic model of what is to ascribe a propositional attitude (see Apperly, 2011). According to this model, ascribing a propositional attitude is to describe or convey a psychological state. Paradigmatically, a relation (attitude) connecting an agent (I, you, she, he) with a proposition (e.g., it is raining; Berlin is the capital of Germany; the building is on fire). The primary goal of a mindreader is to reason about the role of these attitudes in bringing out a particular behavior. For instance, if I attribute to someone the belief that the building is on fire, I can predict her next action will be to run away from the building. Thus, engaging in such reasoning requires the mindreader to describe these psychological states causing behavior.

Before I analyze how this intuition is incarnated in the different analyses

of mindreading, let me provide some definitions to characterize the main ideas. This conception of propositional attitude ascriptions is a type of descriptivism. Descriptivism is a general semantic conception according to which the function of declarative sentences is to state facts, and sub-sentential expressions (names and predicates) denote, refer or stand for objects, properties and relations in the world. This semantic conception is behind what Austin (1962) called the 'descriptive fallacy' (see also Ryle, 1949, 56,115; Ryle, 1979, 89; Belnap, 1990, 1). This fallacy is captured by what nowadays Chrisman (2007) calls "the dogma of descriptivism in philosophical semantics, whereby it's assumed that since semantic content of indicative sentences is standardly given in terms of their truth-conditions, the characteristic function of all indicative sentences is to describe worldly objects, properties, and relations" (p. 227). This semantic dogma, I believe, underlies the propositional attitude ascription models behind the received view in folk psychology.

When the descriptivist dogma is applied to folk psychology, we have a conception of mental state ascriptions as descriptions or (meta)-representations of psychological states. For our purpose, they are descriptions that depict the real psychological states of the subject of attribution.

> **Folk Psychological Descriptivism**: Propositional attitude ascriptions describe or stand for particular psychological entities (beliefs, desires, hopes) causally connected to behaviors and perceptions. Paradigmatically, propositional attitude ascriptions describe an attitude-relation (denoted by the psychological verb) which connects an agent (denoted by the name of the attributee or pronouns) to a content (denoted by the that-clause).

The idea behind folk psychological descriptivism appeared explicitly mentioned in the literature under different labels. For instance, 'postulationism' (Gauker, 2003) or 'factualism'(Bermudez, 2005; Boghossian, 1990)[1]. However, the idea is

[1]The reason to choose the concept FP-descriptivism instead of postulationism has to do with the instrumentalist/fictionalist connotations that the term postulates has. On the one

the same: mental state ascriptions denote the unobservable mental entities of a subject that cause her behavior.

A possible objection to this characterization is that descriptivism is a semantic conception of natural language expressions, while the outputs of mindreading mechanisms are mental states. However, folk psychological descriptivism applies to different semantic bearers. In fact, it is not unusual to find expressions such as 'state-descriptions' (Botterill, 1996) or 'descriptions of psychological states' (Leslie, 2000) as referring to the outputs of mindreading mechanisms. In any case, the importance of characterizing descriptivism is to provide an approximation to a particular understanding of the function that ascriptions and mental concepts play in our social interactions. In this sense, the interest of the definition is to distinguish functionally descriptive-states or sentences from action-guiding states or sentences which do not carry information about the world, i.e., states or sentences which give action-guiding information (Charlow, 2014, 2015; Lewis, 1979; **?**)that specifies particular behavioral patterns. Following Charlow (2014), we can say that descriptive-states encodes a locational perspective, that is, it helps the subject to locate himself in a logical space of possibilities by providing information about the environment. Instead, action-guiding information is a motivational information that helps the subjects to guide his behavior.

3 Descriptivism and Theories of Theories of Mind

The descriptivist dogma behind the orthodoxy is particularly evident among the defenders of theory-theory in both of its versions (modular and theory-formation). For instance, in Gopnik and Meltzoff's (1997) discussions concerning the parallelisms between scientific practices and children's development of theory of mind, we can find several references to the practice of theorizing as

hand, the term 'factualism' opposes 'non-factualism' which could refer to different views that can also be descriptivist, for instance, eliminativism (see next section).

one of finding descriptions or pictures of the phenomenon:

> We might think of childhood as a period when many of the requirements for survival are suspended so that children can concentrate on acquiring a veridical picture of the particular physical and social world in which they find themselves (Gopnik and Meltzoff, 1997, 16)
> Once, as children, we have engaged in the theorizing necessary to specify the features of our world, most of us most of the time may simple go on to the central evolutionary business of feeding and reproducing. But, we suggest, these powerful theorizing abilities continue to allow all of us some of the time and some of us, namely professional scientists, much of the time to continue to discover more and more about the world around us (Gopnik and Meltzoff, 1997, 20).

The primary function of our cognitive development, including folk psychological development, is that 'it gives us a better understanding of the world outside ourselves'. Like scientific theories, the aim of our folk psychological skills is to produce veridical pictures that match up to the outside world, in the mentalizing case, accurate descriptions of the agent's attitude to a particular content.

The same references to folk psychological capacity as picturing unobservable causal events appear in the defenses of the modular version of the theory-theory. For instance, Leslie (1994, 1987) regards theory of mind as involving metarepresentations, involving states that represent others' internal states. Leslie distinguishes between primary representations and decoupled representations or metarepresentation[2]. For Leslie (1987) "the basic evolutionary and ecological point of internal representation must be to represent aspects of the world in an accurate, faithful, and literal way, in so far as this is possible for a given organism" (p. 414). However, what distinguishes meta-representations from primary representations is that in meta-representations, the terms in the that-clause do

[2]A widespread version of FP-descriptivism considers propositional attitude ascriptions as metarepresentations. The term meta-representation is sometimes unclear. The underlying idea seems to be that we have a meta-representation as far as the attributed states are representations. However, we are constantly engaging with representations (reading sentences/looking at traffic signs) in contexts we would not label as meta-representational. The key seems to be that we need to understand mental states as representing or misrepresenting the world in order to understand the role these representations play in producing behavior. In any case, my concern with metarepresentational theories has to do with their descriptive character.

not refer to entities in the world, rather they refer to parts of primary representations in the agent's mind: "Decoupled expressions do not refer to objects, then, they are anchored to parts of primary representations" (Leslie, 1987, 418). Thus, propositional attitude ascriptions describe an agent that has an 'informational' relation to a representation anchored to the primary representation in the agent. In summary, propositional attitude ascriptions are descriptive states that provide information about the internal reality of the target. In this respect, the modularity interpretation does not differ from the theory-formation interpretation. As Gopnik (1996) puts it: "both modularity and the theory-formation theory imply that there is nothing particularly privileged about the representations of folk psychology; this is why modularity theorists can be said to adopt 'the theory-theory' epistemologically... Both these views see folk psychology as a genuinely cognitive phenomenon, a set of representations that are about something in the world" (p. 180). The outputs of the mindreading mechanism, modular or theory-formed, are representations of the psychological states of the target.

Descriptivism is not the sole domain of theory-theory: as Goldman admits, simulation theory is, in the first instance, an approach to mental state attribution, "a species of metarepresentation, an activity in which mental states (beliefs) represent other mental states" (Shanton and Goldman, 2010, 527). Simulation requires projecting mental contents to others; representing others as being in a mental state. In this sense, simulationist theory does not differ from theory-theory or modularist theory:

> On other topics, too, [Simulation Theory] and [Modular Theory] might not conflict. A non-radical version of ST, the kind I favor, cheerfully grants that mindreading involves meta- representations, that is, descriptive representations of mental states. Certainly, a mental attribution is itself a metarepresentational state. So the metarepresentational emphasis of [Modular Theory] is something ST can embrace. This is to agree with Leslie and German's comment that "simulation needs meta-representation" (1995: 133), at least

> where "metarepresentation" is not understood to involve any particular architectural implications.(Goldman, 2006, 100).

Folk psychology requires gaining access to non-observable events. Thus, it demands to formulate descriptions of these events for generating predictions and explanations of the causal consequences of these events (behavior). Summing up, all contenders take as uncontroversial the assumption that mindreading requires descriptions of others' mental states mediating between perceptual inputs and behavioral outputs.

Another way of noticing the pervasiveness of folk psychological descriptivism is through the philosophical discussion of the ontological status of folk psychology. During the eighties, Paul and Patricia Churchland presented several significant arguments to favor *Eliminative Materialism*, according to which our common-sense understanding of the mind is deeply mistaken (P.M. Churchland, 1981, 1984; P.S. Churchland, 1986). As P. M. Churchland puts it:

> The FP of the Greeks is essentially the FP we use today, and we are negligibly better at explaining human behavior in its terms than was Sophocles. This is a very long period of stagnation and infertility for any theory to display, especially when faced with such an enormous backlog of anomalies and mysteries in its own explanatory domain. Perfect theories, perhaps, have no need to evolve. But FP is profoundly imperfect. Its failure to develop its resources and extend its range of success is therefore darkly curious, and one must query the integrity of its basic categories. To use Imre Lakatos' terms, FP is a stagnant or degenerating research program, and has been for millennia (Churchland, 1981, 74-75)

From this view, folk psychology is seen as an unsuccessful scientific enterprise. Our folk psychological vocabulary refers to theoretical entities postulated to understand each other, but neurosciences and neuropsychology prove them to be wrong. In other words, neurosciences will not find a match up between our folk psychological concepts and brain states or processes because "as the eliminative materialists see it, the one-to-one match-ups will not be found, and our common-sense psychological framework will not enjoy an inter-theoretic reduction, because our common-sense psychological framework is a false and

radically misleading conception of the causes of human behavior and the nature of cognitive activity" (Churchland, 1984, 43)

Eliminative materialism contrasts with two views. On the one hand, intentional realism (Fodor, 1975, 1985) holds that our folk psychological vocabulary refers to real functional states implemented in the brain. From this view, our folk psychological concepts reflect more or less accurate descriptions of the mental states and processes underlying our behavior. On the other hand, one may consider that our common sense conception of the mind is a useful instrument for prediction and explanation without assuming ontological commitments about the nature of the states our mental concepts refer. This is the position usually ascribed to Dennett (1987, 1991b). According to (Dennett, 1987), "Folk psychology, then, is idealized in that it produces its predictions and explanations by calculating in a normative system; it predicts what we will believe, desire, and do, by determining what we ought to believe, desire, and do" (p. 52). Rather than theorizing about the internal cause of an agent, folk psychology is a strategy to approach behavior based on a simple heuristic: consider what are the most rational things to believe and desire according to the situation the agent is embedded in, then calculate what is the most rational thing to do according to those stipulations. The rational idealization is neutral about the internal architecture of the agent, and it is not a conjecture about the inner states that provoke the agent's behavior. Dennett's position is subject to different interpretations. On the one hand, it can be considered as sort of fictionalism or 'as if' instrumentalism according to which desires and belief attributions are useful tools for prediction and explanation counted as if they were real entities (McCulloch, 1990; Hutto, 2013). On the other hand, Dennett (1991b) has defended that beliefs and desires refer to 'real patterns' analogous to centers of gravity that, although objective 'fall short of perfection'.

In any case, notice that eliminativism and, at least, some of the interpretations of Dennett's view are still committed to the descriptivist assumption. On the one hand, as Boghossian (1990) argues, the only way to make eliminativism intelligible is by considering it as an error theory (Mackie, 1977). An error theory is a theory that takes a fragment of discourse, mental state predicates in this case, to be empty, that is, nothing in the world possesses the property denoted by mental concepts. As Boghossian puts it: "predicates denote properties and (hence) declarative sentences express genuine predicative judgments, equipped with truth conditions. However, the error theorist continues, because nothing actually exemplifies the properties so denoted, all the fragment's (atomic) declarative sentences are systematically false" (p. 159). Thus, the eliminativist must follow, we should abandon the discourse in question. The key point is: even if you embrace eliminativism, the function of propositional attitude ascriptions is the same, describing or representing the internal causes of behavior. On the other hand, Dennett's view seems to be understood either 'as if' mental state attribution would denote internal entities, or as if it would describe abstract objective entities. But again, this supposes that our ascriptions are used to describe entities (abstract or fictional).

As a result, the three ontological approaches share the view that the accuracy of human theory of mind, like scientific theories, should rely on the existence of the worldly counterpart of their concepts. They share the idea that the primary function of declarative sentences in general, and ascriptions in particular, is to describe or state facts fixed as truth values. Furthermore, all scholars in the business of folk psychology seem to share the starting point of considering folk psychology as an epistemic enterprise. The function of folk psychology is to gain knowledge of causal mechanisms bringing out behavior for the sake of explanation and prediction.

4 FP-Descriptivism, Language and Mentalizing

The strategy I pursuit in this dissertation is to explore the consequences of a particular treatment of propositional attitude ascriptions for the debate concerning the role of language in mentalizing. This treatment seeks to be an alternative to folk psychological descriptivism. However, this enterprise would be pointless if the different parts in debate concerning the language/folk psychology relation do not assume descriptivism as a critical linchpin of their views. This section aims at demonstrating that FP-descriptivism is central for characterizing the different perspectives reviewed in chapter 2, but also, that it plays a key role in the articulations of the arguments and the interpretation of the evidence supporting each characterization. In order to discuss it, let me consider apart the two groups of theories again.

4.1 The FP-Descriptivism behind the Communicative Views

We can find explicit references to FP-descriptivism in the defenders of both versions of the communicative view. For instance, Perner (1991) argues that children's acquisition of folk psychological capacities requires a conceptual change that involves understanding others' minds as representational entities causally connected to action: "Modern particle physics has given a theory of gravity, that is, an explanation of why celestial (or any other) bodies exert a force on other bodies. My claim is that a representational view of the mind serves a similar role in the child's understanding of mental states and for that reason might be called theory of mind" (p. 124). Then, he concludes that 'mental states are theoretical constructs that are necessary for explaining the behavior of very complex information-processing systems" (p. 109). Furthermore, Perner understands folk psychology in metarepresentational terms that he defines as "the ability to represent that something (another organism) is representing some-

thing" (Perner, 1991, 7). That is, folk psychology understands propositional attitudes as 'internal representations' that other agents have in mind. Analogous support to descriptivist assumptions can be found in the defenders of the strong version (Bloom, 2000; Carruthers, 2013; Fodor, 1978; Leslie, 1994; Origgi and Sperber, 2000; Sperber, 2000). For instance, Fodor (1978) expresses it as follows: "Propositional attitudes should be analyzed as relations. In particular, the verb in a sentence like 'John believes it's raining'" expresses a relation between John and something else, and a token of that sentence is true iff John stands in the belief-relation to that thing. Equivalently, for these purposes, 'it's raining' is a term in 'John believes it's raining'" (p. 501). According to Fodor's view, propositional attitude ascriptions describe a relation between an agent and an object, which in his view, is a sentence in the language of thought.

In principle, the commitment to descriptivism of both versions of the communicative view is orthogonal to the theoretical position concerning the role of language in theory of mind you advocate for. The problem, I believe, is that FP-descriptivism plays a significant role in the argumentative strategies of both versions of the communicative view. Consider again the basic arguments on each side supporting the communicative view. According to the weak version, the conceptual change that provokes the acquisition of full-fledged theory of mind around the fourth year of age does not require language apart from considering it as an informational source. In Perner's version, the conceptual change appears when the 'situation theory' that children acquire before the age of 4 confronts some explanatory problems. Then, children are compelled to come up with a different view to explain those anomalies. According to Perner, the conceptual change is driven by explanatory demands when the child is exposed to certain information her situation theory cannot accommodate. Thus, if language plays any role in this conceptual change, it is just as a source of empirical evidence

that strengthens or undermines the child's theory.

On the other hand, some strong version's argument relies on a particular understanding of communication as an ostensive-inferential process, according to which utterances are only cues to grasp the *intended meaning of the speaker*[3]. In order to grasp the intended meaning, the hearer must infer it from the information provided by the utterance and context. As we saw in the previous chapter, this process is driven by a principle of relevance searching. In order to trigger this process, the speaker must token two intentions: an informative intention that makes manifest to the audience a set of assumptions; and a communicative intention that makes mutually manifest among speaker and hearer that the communicator has the informative intention. As we saw in the previous chapter, the intended meaning and intentions must be understood in descriptivist terms.

My point of contention is that both arguments are not compelling when FP-descriptivism is removed. The two cases start from the idea that we need to form descriptions of the internal reality of the subject we want to interpret subjects. From the weak side, FP-descriptivism supports the intuition that what children must learn is a theory that helps them to produce and derivate consequences from representations of other's mental states. Once you contemplate the possibility that the primary function of folk psychology may not be to describe others' psychological states, then the necessity of forming a theory to bridge the epistemic gap between behavior and perceptions is not required. But also, it opens the possibility that folk psychological abilities would have a different function which requires language to develop the required conceptual skills. Thus, the influence of language as a source of empirical information that sophisticates the theory of mind collapses with the model as well. In a similar vein, the ostensive-inferential model supporting the strong version only works under the assumption that the speaker's meaning is an unobservable mental

[3]The other way to support the cognitive view was to maintain that mindreading capacities are innate on the basis of implicit FB-tasks. I will discuss the meaning of these experiments for my view in chapter 6.

state we need to access. In other words, the propositional attitudes required in communication are unobservable entities we can only access through an inferential process. Again, once descriptivism is removed, the ostensive-inferential model is not necessary, because we do not need to understand the intended meaning in terms of internal representations we cannot access without inferential capacities. The two versions of the communicative views support what Gauker (1995, 2003) dubs the Lockean theory of language according to which:

> The central function of language is to enable a speaker to reveal his or her thoughts to a hearer. The speaker has a certain thought in mind and intends the hearer to recognize that he or she has that thought in mind. The speaker chooses his or her words in the expectation that on the basis of the words spoken and the circumstances of utterance, the hearer will be able to infer that the speaker has that thought (Gauker, 2003, 3).

This perspective of language reinforces FP-descriptivism. It is based on the idea that the meaning of our communicative actions must be understood in terms of the meaning of internal mental states we need to meta-represent. In this sense, the outcome of the linguistic processes is to infer and provide cues for engaging into descriptions of others' propositional attitudes.

4.2 The FP-Descriptivism behind the Cognitive Views

The clearest appeal to descriptivism in the cognitive side comes from the semantic and the second-order dynamics views. According to the semantic view (Olson, 1988; Segal, 1998; Smith, 1996), our folk psychological capacities are bound to the linguistic capacity of using verbs such as want, believe or think in conversations. These concepts facilitate the acquisition of metarepresentational skills necessary for engaging in folk psychological explanation and predictions. For instance, Segal defends a form of descriptivism when saying:

> There are three basic kinds of inferences that the psychology faculty makes about propositional attitudes: influence of the world on the attitudes, intra-mental interactions, and influence of the attitudes on action. In each case, what is crucial about the attitudes is their

> representational properties. In representing what someone believes or desires, one needs to represent the representational properties of the attitude in question. And a natural way to do this would be to represent the individual standing in a relation to a representation of some kind - something like a proposition, or a Fregean Thought, or a total form. Total forms are representations. So representing attitudes as relations to total forms serves as way to represent the attitude's representational properties (Segal, 1998, 153).

Folk psychology (psychology faculty) requires descriptions of internal psychological states in order to deal with the intra-mental interactions mediating between the previous influence of the world and the subsequent behavior. Again, the emphasis is the necessity of descriptive-states that represent the relations between agents and internal states.

In contrast to the semantic view, the second-order dynamics view bounds the descriptions of psychological states to linguistics abilities broadly considered. It is language as a representational vehicle, rather than specific concepts, which is required to engage in metarepresentations (Bermudez, 2003a,b, 2005). Bermudez reveals his descriptivist assumptions as follows:

> The key point is that propositional attitude mindreading does not involve thinking about a direct relation between a subject and their environment in the way that perceptual mindreading does. That is the whole point of the Sally-Anne task and the various other false belief tests. Beliefs can be false and false beliefs are just as powerful in bringing about behavior as true ones. It is what subjects believe about the world that explains and predicts their behavior. This means that representing another subject's belief state requires representing them as having representations of their environment – representations that can be either true or false. Philosophers typically analyze belief (and other propositional attitudes) as an attitude to a proposition or thought. The terminology is inessential, however. What matters is that understanding what another subject believes requires metarepresentation in a way that understanding what they are seeing or hearing does not (Bermudez, forthcoming, 379)

Bermudez points out that representing a propositional attitude requires understanding other persons as representing the environment, that is, as possessing an inner state with representational content. Again, mindreading is understood as a process that requires to describe the psychological reality of the target. Notice

that the kind of function that metarepresentational states play in Bermudez's view does not differ from other descriptions. According to him, when we say that 'Anne believes that the ball is in the box' we are stating a fact, we are not doing something different from stating, for instance, that 'the laptop is on the table'. The difference is that in propositional attitude ascriptions one of the objects (the proposition) is semantically evaluable. However, the functions of both states are the same, namely, describing a particular state of affairs.

The arguments of both perspectives rely on FP-descriptivism. The main motive behind both the semantic and the second-order dynamics view is the idea that language paves the way to acquire certain inferential capacities that otherwise would not be available to the subject. These inferential capacities involving metarepresentations are a requirement to predict and explain mental entities behind others' behavior. In other words, reasoning involving metarepresentations is required because folk psychology is a question of revealing internal states of the subject. In this sense, the arguments behind both the metarepresentational and the semantic views are based on descriptivist assumptions. Thus, if FP-descriptivism turns to be false, the arguments supporting both views would lose its strength.

In contrasts to the aforementioned cognitive views, the syntactic approach holds a different connection to FP-descriptivism. Notice that the main motivation behind the other views (the semantic and second-order dynamics views) for supporting a particular understanding of the relation between language and mindreading is purely theoretical. Although they have empirical aspirations, the main arguments favoring the views rely on conceptual and theoretical connections between language and mindreading. For instance, Perner's conception of the relation seems to be motivated by his commitments to the theory formation view of mindreading development; or for instance, Bermudez's view is favored

by a particular understanding of the representational capacities of language. On the contrary, the motivation behind the syntactic view is mostly empirical. It rest on empirical data that putatively demonstrate causal connections between language and mindreading. Thus, although it supports a descriptivist view of propositional attitude ascriptions, the attack to FP-descriptivism does not necessary undermine the thesis. As a result, any language/mindreading thesis based on a non-descriptivist conception of mental ascriptions should account for the empirical facts that motivate the syntactic view.

De Villiers (2007)claims that the key of syntactical complementation contribution to theory of mind relies on the fact that "Even if the proposition in the lower clause is false –there was no wasp – the whole sentence remains true" (p. 1867). Furthermore, he implicitly claims that these syntactic capacities open up the possibility of accessing mental entities. This is shown when defining folk psychological capacities: "Theory of Mind refers to the folk psychological theory that we use to predict and explain others' behavior on the basis of their internal workings: their feelings, intentions, desires, attitudes, beliefs, knowledge, and point of view. That is, we need to posit a mental state inside a person to accommodate the occasional disjunction between an external stimulus and a response" (p. 1859). Notice that the syntactic view is subject to two interpretations. Firstly, it could be interpreted as claiming that complementation is a requirement for passing the FB-task. Secondly, according to a more substantial interpretation, complementation is a requirement for passing the FB-task because it provides the necessary structures to form the mental states that stand for the psychological states of the attributee. Notice that the correlation and training studies reviewed in Chapter 2 only support the former interpretation, but de Villiers seems to support the latter one. The latter interpretation seems to be committed to FP-descriptivism, and thus, if FP-descriptivism turns to be

inaccurate, de Villiers's interpretation would be undermined.

Summing up, although all approaches to the influence of language in social cognition usually take FP-descriptivism for granted, the role it plays in the arguments is unequal. It is fair to conclude that the weak and strong versions of the cognitive view and the semantic and second-order dynamics views give to FP-descriptivism a central role in their arguments. Thus, if we have reasons to motivate an alternative to FP-descriptivism, those reasons would undermine the strength of the argumentative strategies behind these views. Furthermore, although the syntactic view can be undressed of descriptivist assumptions, their supporters seem to be committed to the descriptivist interpretation of the empirical evidence concerning syntax. Thus, they are exposed to any argument against it.

5 Conclusions

In this chapter, I made explicit a basic philosophical assumption behind the different hypotheses concerning the influence of language in folk psychology: Folk Psychological Descriptivism. This assumption relies on the idea that our propositional attitude ascriptions describe the mental entities mediating perceptual inputs and behavioral outputs of the targets. Folk psychology aims to reason about these states in order to predict and explain behavior to carry out different social interactions. This idea is quite pervasive in developmental psychology and other areas of research issuing the impact of language in folk psychological competences. My central objective was to expose how the central argumentative strategies behind the approaches to language/theory of mind interface gravitate toward FP-descriptivism. The main conclusion of this chapter is that if FP-Descriptivism turns to be wrong, most of the above-mentioned theories concerning the language/folk psychology relation would see their basic argu-

mentative strategy jeopardized. This conclusion justifies the indirect strategy I advocate for in this dissertation, namely, providing an alternative to descriptivism from which we can rethink the problem of the influence of language in social cognition.

Chapter 4

Against Folk Psychological Descriptivism

1 Introduction

The previous chapters were mainly devoted to characterizing the received view of propositional attitude ascriptions. I argued that most theories concerning the language/mindreading relation share a basic assumption: Propositional attitude ascriptions describe or stand for particular psychological entities (beliefs, desires) causally connected to behaviors and perceptions. This assumption is motivated by the idea that folk psychological practices constitute an epistemic enterprise consisting of accessing the psychological states for the sake of explanation and prediction. For instance, ascriptions of the desire to have a beer and the belief that there is a pack of ales in the fridge is explanatory of Peter's action of opening the fridge because these ascriptions refer to the inner states of Peter that caused this action. Although this picture, I admit, is highly persuasive, it faces some serious problems.

The aim of this chapter is to provide arguments defending that the attributions of beliefs and desires should not be understood as descriptions of the inner psychological machinery producing behavior. Firstly, I present some uses of propositional ascriptions which have been obviated by the received view. Arguably, these uses are not only important to navigate our social world, but also, they are not easily interpretable under FP-descriptivism (Section 2). Secondly, I present an argument concerning disagreements in social circumstances. Folk psychologists usually disagree about how to interpret a social situation. I argue that some of these disagreements present a serious challenge to FP-descriptivism (Section 3). Finally, I argue that explanations in terms of beliefs and desires are normally framed into contexts of accounting for counter-normative behavior (Section 4). If this is right, ascriptions involve an evaluative element that FP-descriptivism cannot account for.

2 The Diverse Nature of Ascriptions

In the previous chapters, I defended that most of the positions in social cognition have embraced the same philosophical view concerning propositional attitude ascriptions: FP-descriptivism. Part of the motivation behind it is that FP-descriptivism intuitively accounts for explanation and prediction: propositional attitude ascriptions are explanatory/anticipatory of behavior because they describe its causes. Consider these explanations:

(1) Marta came late because she thought that the meeting was at eight o'clock

(2) John went for coke because he doesn't want to drink beer tonight.

By describing the inner machinery of John and Marta, we can explain and predict their behavior. Even though the persuasiveness of the picture is indisputable, the received view has tended to exaggerate the explanatory/predictive

function of propositional attitude ascriptions (Andrews, 2012; Gauker, 2003; Morton, 1996). This exaggeration has made the received view overlook some important uses of propositional attitude ascriptions. This section has two purposes. Firstly, it aims to argue that some uses of propositional attitude ascriptions do not have an obvious interpretation from descriptivism. That is, certain uses of propositional attitude ascriptions are not compatible with the idea that ascriptions describe internal states of an agent. Secondly, these uses reveal an evaluative function of propositional attitude ascriptions which cannot be explained in descriptive terms.

2.1 Parenthetical uses

The philosopher of language JamesUrmson (1952) introduced parenthetical verbs to characterize verbs in the first-person present which can occur idiomatically in any of these forms:

(3) I suppose (that) your house is very old.
(4) Your house is, I suppose, very old.
(5) Your house is very old, I suppose.

According to Urmson, the verbs which can occur in these forms are not psychological descriptions. Urmson contrasts them with descriptive predicates, for instance, 'miserable', which are not idiomatic in these forms:

(6) I am miserable because he is unwell
(7) *He is, I am miserable, unwell

The fact that the utterance (7) is not idiomatic shows, according to Urmson, that 'I am miserable' has a descriptive function. Besides this grammatical feature, Urmson characterizes parenthetical verbs for their pragmatic function: "to indicate the evidential situation in which the statement is made (though not to

describe that situation), and hence to signal what degree of reliability is claimed for, and should be accorded to, the statement to which they are conjoined" (p. 485). By using a parenthetical verb, the speaker indicates how the content of the proposition fits logically, evidentially and emotionally into the context.

Parenthetical verbs have also received considerable attention in linguistics (Aijmer, 1997; Goddard, 2003; Thompson and Mulac, 1991; Wierzbicka, 2006). In fact, there is a widely accepted conception of parenthetical use of mental concepts as non-descriptive. For instance, the phrase 'I think' is frequently presented as having the function to mitigate the degree of commitment to the sentence it ranges. Wierzbicka (2006) provides a deep analysis of parenthetical uses of 'believe', 'think' and other mental verbs. She claims that the verb 'think' conveys the meaning of disclaiming knowledge "not by saying "I don't know" but by saying "I don't say: I know it." (p. 38). In other words, 'I think P' expresses a certain degree of caution. Similarly, the verb 'believe' (in contrast to 'I think' for instance) seems to play an indicative function. As Aijmer (1997) claims: "I believe does not only express a subjective attitude. It also conveys that the speaker has some evidence for what he says." (p. 17). We can see the contrast between 'I think' and 'I believe' in the incompatibility of 'I believe' with phrases like 'I'm not sure'. While 'I think that Riga is the capital of Latvia, but I'm not sure' is idiomatic, 'I believe that Riga is the capital of Latvia but I'm not sure' is not. This difference between the level of reliability that 'think' and 'believe' convey must not divert our attention away from the fact they share their basic function: they are devices for canceling or altering the speaker's commitments. The verbs 'believe' and 'think' seem to be mitigators of the force of the claim. Of course, parenthetical uses are not restricted to these types of indications involving mitigations. Verbs as 'rejoice' or 'regret' indicate emotional orientation, others as 'wish' or 'desire' indicate the preference toward

the content of the statement. What these parenthetical uses of propositional attitude verbs share is its function for providing indications or prescriptions to the hearer about how to evaluate the statement.

Another interesting point concerning parenthetical uses of these verbs is that they appear quite frequently. For instance, Thompson and Mulac (1991) found that 'I think P' has an enormously higher frequency than 'I think that P' in spoken language, which they take as a mark of the parenthetical use of 'think'. Conforming to the data they handle, 'I think that P' occurs 122 times (7% of the total I think), whereas 'I think P' occurs 1,644 times (93%). These figures call our attention upon the fact that these uses of mental predicates play a significant role in our practices. Although they rarely appear in the analysis of propositional attitude ascriptions or mindreading debates, it seems they are quite pervasive in our social practices. In a similar line, Scheibman writes:

> Eighty-nine percent, then, of all present tense 1s + verb of cognition combinations are formulaic expressions such as I think, I don't know, and I guess, and these expressions function epistemically or serve to mitigate assertion or disagreement in conversation. These are subjective functions that organize expression of the speaker's point of view in conversation; they are not being used to inform participants of the speaker's cognitive activities" (Scheibman, 2001, 70-71).

This function of parenthetical uses is quite common in our everyday life. Hence, given that FP-descriptivism attempts to account for belief and desire ascriptions in our social interactions, it must be able to explain it.

Now, it is important to see the difficulty of analyzing parenthetical uses under a descriptivist reading. Consider the following utterances:

(8) Ronda Rousey is the most dominant sportsperson in the world [said by Víctor]

(9) Ronda Rousey is, I believe, the most dominant sportsperson in the world [said by Víctor]

According to FP-descriptivism, (9) refers to internal states of a subject. Utterance (9) states that a person (Víctor) is in a certain relation to a psychological

state or content (that Ronda Rousey is the most dominant sportsperson in the world). On the contrary, (8) states that certain person (Ronda Rousey) possesses a particular property (being the most dominant sportsperson in the world). From this reading, the two sentences communicate different states of affairs. However, this seems to be counterintuitive since (8) and (9) seem to express something quite similar. They can be asserted in the same circumstances without changing the contribution to the conversation substantially. In fact, the only difference seems to be, as Wierzbicka and Urmson claim, that the speaker indicates with (9) that the statement must be taken with caution.

One may question the importance of the parenthetical uses for social interactions aside from its obvious linguistic role. Contrariwise, possessing conceptual tools to indicate, for instance, different degrees of reliability is useful to avoid public sanctions in cooperative projects. Several studies in empirical economy, anthropology, and evolutionary biology demonstrate the importance of punishment and trustfulness in cooperative behavior. Humans exhibit a strong tendency to punish non-cooperative behavior[1] (Fehr and Gächter, 2002; Henrich, 2004; Henrich et al., 2006; Richerson and Boyd, 2008). These studies confirm the existence of a human tendency to punish in spite of the cost, and the efficacy of maintaining cooperative behaviors. A possible explanation of this tendency could be the advantage of reputation (Nowak and Sigmund, 1998; dos Santos et al., 2010; Wedekind and Milinski, 2000). Punishing defectors and cooperating

[1]The empirical evidence is obtained through two different paradigms: the ultimatum game and third-party punishment game. In the ultimatum game, player 1 receives an amount of money. She can offer a portion to player 2. Player 2 can accept this portion or reject it. In principle, a self-interested player should accept any portion distinct from zero since rejection leaves both players without money. However, this rarely happens in any culture. In the study conducted by Henrich et al. (2006), 1762 subjects from 15 countries exhibit the same tendency to punish player 1 when he is 'unfair'. The same results are found in the third-party punishment game. In this game, there is a player 3 who receives half of the money assigned to player 1. Player 1 can choose a portion to offer player 2. Now, player 2 has no power of decision. However, player 3 can choose between paying 10% to punish player 1 to suffer a deduction of 30%. In principle, player 3 should never choose to punish player 1 because of the costs; nevertheless, similar patterns to the previous games were found.

with punishers emerges as a dominant strategy where the reputation is the key driving the evolution of cooperation. In a context where reputation and the punishment of non-cooperative behavior are important, it makes sense to develop conceptual tools that allow agents to indicate the degrees of reliability of their assertions. Imagine a cooperative context where several agents are trying to perform a common goal, for instance, hunting a deer. If an agent claims 'the deer is by the river', he is open to sanction from the other cooperators if the statement turns to be false. However, indicating that the content should be taken with certain caution can avoid possible sanctions in those cases. Furthermore, these indications are useful in cooperative decision-making. In contexts where a group of people needs to make some joint decision, it is necessary to throw different hypotheses or state different facts concerning the decision. Again, indicating different manners to take these contents to your peers can be useful. Imagine a group of friends deciding how to surprise a mutual friend (Noel) with a party. One could say: 'Noel comes back from work at six o'clock'. Considerations about the level of reliability or preference are at stake in order to make the decision. Parenthetical uses of propositional attitude verbs are important in our social contexts. We use them to indicate to our peers how they must evaluate our assertions.

2.2 Communicative uses

In the previous subsection, I have introduced parenthetical uses as examples of propositional attitude ascriptions which cannot be easily accommodated under a descriptive reading. However, those uses do not raise a problem for FP-descriptivism as far as the paradigmatic examples of folk psychological explanation are third-personal. The clearest examples are those used in FB-task, where a child says 'Sally believes that the doll is in a box A' and the ascriptions

seems to refer to the mental states of Sally. The aim of this subsection is to start to glimpse a non-descriptive reading of third-person ascriptions analogous to those introduced by parenthetical uses. In order to do that, consider what is called the communicative uses of propositional attitude ascriptions. These uses of belief and desire ascriptions are introduced by Christopher Gauker (Gauker, 2003; Cleave and Gauker, 2010) to exemplify his approach to belief and desire ascriptions. Consider these examples used in Cleave and Gauker (2010, 310):

> Scenario One: Billy and Sally are playing with their toys in the living room. Mother is expecting guests in half an hour. Sally steps into the kitchen where her mother is preparing food. Mother says to Sally, 'You and Billy, please pick up all your toys in the living room and take them to your own rooms'. Sally returns to the living room and says to Billy, 'Mom wants us to put away the toys'.
>
> Scenario Two: Billy and Sally are playing in the yard. They expect to attend a friend's birthday party later in the afternoon, but they are not sure when it starts. They see another friend, Markie, at the end of the block. Sally walks down the street to ask when the party starts. When she returns, she says to Billy, 'Markie thinks the party starts at four o'clock'.

These scenarios exemplify uses of belief and desire ascriptions that are quite peculiar. In scenario one, Sally is giving a command on her mother's behalf. The function of the desire attribution seems to be conversational, that is, Sally tries to give a command to her brother on her mother's behalf. Similarly, Sally is asserting something on Markie's behalf in scenario two. Sally is conveying certain information to Billy as being said by Markie. Gauker argues that these uses are the primary practice of making attributions. In fact, he argues that explanatory uses of ascriptions derivate from communicative uses[2].

[2]Acero and Villanueva Fernández (2012) offer compelling arguments in favor of the idea that Wittgenstein would have a similar position in mind. This interpretation goes against the classical interpretation of Wittgenstein's work, according to which he maintained an expressivist position about first-person ascriptions (Avowals), while defending a descriptivist position about third-person ascriptions (see also (see also Villanueva Fernández, 2017a).

Leaving aside Gauker's communicative view[3], the communicative scenarios are relevant for our purposes because they present the same problem to descriptivism we saw in the case of parenthetical uses. Arguably, the contribution to a conversation made by 'Markie thinks the party starts at four o'clock' is quite similar to Markie's assertion 'the party starts at four o'clock'. Similarly, the attribution of Sally plays the same role on the conversation that it would have played a mother's command. Now, the question is which functions the propositional attitude verbs have in this kind of uses. Cleave and Gauker (2010) present two possible candidates: voicing authority and corroboration/disclaiming. Firstly, as in Scenario one, we regularly talk on others' behalf in order to exercise their authority. Sally is exercising her mother's authority when claiming she wants them to pick up the toys. Secondly, as in Scenario two, we talk on others' behalf in order to assert something we are not in a position to assert or we think others are in a better position to assert. A possible reply at this point would maintain that it is the description of the appropriate mental state, for instance, the mother's mental state, which is playing the role of exercising the authority. We must notice that Sally is doing something different than stating a fact. Consider a different example, when my friend Antonio promises me to go to the cinema ('I promise you to go to the cinema'), I can describe the event by saying 'Antonio promises me to go to the cinema'. By reporting his promises, I am not doing anything like promising on his behalf; I am reporting what he did. However, Sally is doing something similar to commanding when reporting her mother's desire. The peculiarity of the use of 'want' in this context is that it makes the ascriptions work as a command. In

[3]The positive view about propositional attitude ascriptions I canvass in this dissertation (next chapter) is compatible with the communicative view and its intuitions. In spite of it, I will present my position under a different frame. However, I believe that both positions are not incompatible. They both share the same points: (1) the refusal of descriptivism and (2) the necessity of accounting for the evaluative character of belief and desire ascription (see section 4, this chapter).

this sense, the verb 'want' in the Sally case works more as the verb 'promise' said by Antonio when making me a promise than when I use it to report his action. 'Mom wants us to put away the toys' is not stating a fact.

An important point to notice here is that communicative uses seem to exhibit a function similar to this of parenthetical uses. Vicarious speech acts make sense when we want our hearer to evaluate the speech act as being performed by another person. Once we acquire these communicative devices (belief and desire ascriptions), we can talk on the other's behalf, even when they didn't perform such a speech act:

> if two people are present to one another, then they can converse by making assertions and commands. But even if someone is absent from a conversation, he or she can, to a certain extent, participate in the conversation inasmuch as he or she may be represented by someone present who makes assertions and commands on his or her behalf (Gauker, 2003, 226)

Interestingly, Gauker argues that this capacity leads us to acquire the capacity for predicting and explaining others' behaviors. He defines the capacity of explaining and predicting in terms of inner assertions and commands (see also Tooming 2016, forthcoming for a defense and amend of the communicative view[4]). Aside from considering whether or not the communicative uses can account for explanatory/predictive functions, the relevance of these uses lies in revealing that propositional attitude verbs have a function which is not descriptive: exercising authority, credibility or uncertainty. This function, I shall argue in section 4, can account for the explanatory uses of propositional attitude ascriptions as well.

[4]Tooming (forthcoming) takes the justificatory position I use to frame my view in the next sections as a contender against both descriptivism and the communicative view. However, I consider my evaluative position as a way to make some intuitions of both the justificatory view and the communicative view compatible.

2.3 The Problem of the Diversity of Uses

In this section, I have presented two different uses that propositional attitude ascriptions play in social interactions. These functions share a basic feature: they are not readily interpretable under the FP-descriptivist view. Parenthetical and communicative uses seem to have an evaluative function: they indicate a particular assessment with regard to a certain parameter (credibility, authority, certainty...). The two uses exhibit a particular feature of propositional attitude ascriptions that can be defined as an evaluative attitude toward the attributee, burden her with certain credit, significance, merit or demerit, and so on. When introducing these verbs, the speakers indicate certain degree of responsibility of the attributee toward a content that is manifested in attitudes of approval or disapproval. Of course, propositional attitude verbs are pretty diverse and they can indicate different evaluative attitudes. Verbs as 'believe' and 'know' introduce different degrees of responsibility or merit. Other verbs, such as 'hope' and 'guess', are more related to conviction. 'Desire' and 'want' serve to express certain neutrality, disclaiming or particularization (see chapter 5). However, they all share the practical function of regulating and adjusting actions. Strandberg (2012) characterizes this feature of evaluative concepts by appealing to the practicality principle. According to this principle, evaluative language presumes an action-guiding attitude, i.e., it expresses a special connection with action. Evaluative language is practical because "it enables us to regulate one another's behaviour... (it) is generally utilized to get us to act, or to get us to refrain from acting, in certain ways so as to adjust our various actions in relation to one another" (Strandberg, 2012, 89). Notice that the parenthetical expression 'I think' serves to refrain from the actions associated to the sentence because the speaker is indicating she is not fully committed to P. Similarly, third personal attributions make the attributee responsible for

the actions connected to the ascription. This action-guiding component is not captured by descriptive expressions denoting worldly aspects.

In a nutshell, there are some uses of propositional attitude ascriptions which serve to evaluate attributees. These evaluations presume an action-guiding component which regulate behavioral patterns. Certainly, the existence of these uses does not serve as an argument against FP-descriptivism generally considered. However, they reveal some uses of propositional attitude verbs whose meaning is non-descriptive. In other words, some instances of mental terms, such as 'believe' or 'desire', do not denote relations or entities of any kind. Furthermore, as I will argue in the rest of the chapter, paradigmatic cases of propositional attitude ascriptions in folk psychological situations are instances of those evaluative uses. .

A way to resist this argumentative strategy would be to defend that whereas the content of propositional attitude ascriptions is descriptive, the evaluative component emphasized by parenthetical or communicative uses is somehow pragmatically conveyed. The most sophisticated versions of these options would take the evaluative component of propositional attitude ascriptions as being captured by linguistic mechanisms such as presuppositions, conversational implicatures, or conventional implicatures. Several authors (Barker, 2000; Copp, 2009; Finlay, 2004; Potts, 2005; Strandberg, 2012) have defended different versions of this idea to capture the evaluative component of moral terms. In order to make my point, I will briefly present these views, and then several arguments to question them (The particular arguments against the views I present here can be found along others in Buekens 2011; Fletcher 2014). Firstly, one may argue that expressions where 'I think' has a parenthetical use, have a descriptive content of the type 'the subject is in a relation (think) to a propositional content P', but the disclaim of responsibility is somehow presupposed. In other words,

one may argue that the evaluative component of the propositional attitude ascriptions is a presupposition. The problem for this view is that presuppositions generally exhibit a semantic behavior different from the evaluative component. For instance, while presuppositions can be blocked when the expression is embedded in a larger construction, the meaning of parenthetical or communicative constructions cannot(Karttunen, 1973). For instance, if I say 'it was John who solved the problem' there is a presupposition that the problem was solved by someone. However, if the sentence is embedded in a larger expression as "if the problem was solved, it was John who solved it", the presupposition is blocked. This does not happen with the meaning of parenthetical expressions, for instance, in "it was John, I think, who solved the problem", and "if the problem was solved, it was John, I think, who solved it; the evaluative component of 'I think' is not altered in the sentence (see Stanley, 2015, 137)

Secondly, one may argue that the evaluative component is implicated when a Gricean maxim is violated (Grice, 1975). For instance, one may argue that by saying:

(10) Riga is, I think, the capital of Latvia

The speaker is violating a submaxim of quantity ("Do not make your contribution more informative than is required") because assertions express what is believed by the speaker. Thus, by violating this maxim, the speaker is implicating that she is not sure about her assertion. The main problem with treating these uses as cases of conversational implicatures is that conversational implicatures have certain features that the expressive meaning of parenthetical uses does not; for instance, their cancelability. Consider a classical example of conversational implicature:

(11) A: 'What do you think of the new boss?'
B: 'It sure is sunny at this time of year, isn't it!'

[Implicature: the boss is not competent]

In (11), the speaker can cancel the implicature by denying she means that the boss is not competent. However, the speaker cannot deny that she is expressing uncertainty with 'I think' in (10). Similarly, the speaker cannot deny that he is making another person responsible for the claim when using a communicative expression as 'Mom wants us to put away the toys'.

Finally, one may consider the evaluative components of ascriptions as a type of conventional implicature. A conventional implicature can be characterized as follows: P is a conventional implicature of Q if and only if: (1) the meaning of P is encoded in Q, (2) the meaning of P does not affect the truth or falsity of Q, and (3) P is entailed by Q Potts (2005). According to Potts, parenthetical expressions are examples of conventional implicatures:

(12) Lance Armstrong, the doped cyclist, battled cancer.
Descriptive content: Lance Armstrong battled cancer.
Conventional implicature: Lance Armstrong is a doped cyclist

Admittedly, this characterization seems to capture the semantic nature of the evaluative component, that is, this component does not affect the truth conditions of P but nevertheless is part of the conventional meaning. The problem is that when analyzing parenthetical uses of mental verbs as conventional implicatures (Riga, I think, is the capital of Latvia) the implicated meaning would be redundant (I think that Riga is the capital of Latvia). Thus, the parenthetical uses of propositional attitudes do not seem to fit into Pott's theory of conventional implicature.

Leaving the particular problems that these theories have, there is a general objection that can be presented against all these ways to capture the evaluative function of propositional attitude ascriptions. Implicated and presupposed content is usually characterized in propositional terms, i.e., descriptive content

is a proposition. In other words, I can only implicate or presuppose information that can be evaluated as true or false. However, as I argued in chapter 3, evaluative information is action-guiding information that provides specifications about how to behave (Charlow, 2014; Lewis, 1979). This information cannot be captured by propositional content. For instance, when ascribing knowledge to someone the attributer is giving some credit to the attributee by endorsing the truth of the content ascribed. Imagine someone saying 'the door, Edu knows, is closed'. The above mentioned views would characterize the ascriptions as follows:

(13) Descriptive content: Edu is in a mental state with the content the door is close

Implic./pres.: the door is close is true according to the standard S

Notice that if the implicated/presupposed content is made explicit, the evaluative function of the knowledge ascription disappears. In fact, while someone saying (13) could say that he is not committed to support S, someone saying "Edu knows that the door is closed" cannot. The only way to capture the attribution of merit through the support to the standards would be by adding, as something implicated or presupposed, an expression of the form "I support S" and this would imply to introduce a performative verb, which are by definition non-descriptive (Austin, 1962). In a nutshell, evaluative uses and expressions seem to carry some type of action-guiding information which cannot be captured propositionally (for other ways to capture the non-propositionality[5] of the evaluative content see Charlow 2014, 2015; Stalnaker 2014; Lewis 1979). Of

[5]One could insist in a theory of implicature or pressuposition by arguing that the implicated or pressuposed content must be characterized in action-guiding terms. However, this would imply to argue that it is a special type of implicature or pressuposition (Cepollaro and Stojanovic, 2016). Although I am inclined to consider this view would be problematic, there is no need to embrace a particular option at this point. The minimal commitment I subscribe here is that the evaluative component of propositional attitude verbs is part of the conventional meaning and is characterized in terms of action-guiding information; considering one theory or another to model this semantic aspect is beyond the scope of this work.

course, this discussion is far from exhaustive. Probably one could come up with other attempts to capture the evaluative component in pragmatic terms (see Villanueva Fernández, 2017b), but I think the objections I just presented gives a feeling of the problems that such an attempt would have to face.

3 Normative Disagreements and FP-Descriptivism

In the philosophy of mind, several authors are concerned with the problem of the under-determinacy of propositional attitude ascriptions (Davidson, 1970; Dennett, 1987; McCulloch, 1990; Quine, 1960; Slors, forthcoming). In several occasions, our folk psychological ascriptions "fail to yield clear, stable verdicts about which beliefs and desires to attribute to a person" (Dennett, 1987, 29). This under-determinacy is particularly clear in contexts where two interpreters disagree about a particular attribution (Dennett, 1978a; Field, 2009; Pérez-Navarro et al., MS; Spaulding, Venue). This under-determination has been presented as an argument to motivate different approaches to the ontological load of propositional attitude ascriptions (for instance, Dennett's interpretationism) or to cast into doubt the accuracy of the standard theories of mindreading (Spaulding, Venue).

In certain social interactions, disagreeing about others' intentions, beliefs or desires is not unusual. Frequently, those disagreements persist despite the evidence regarding the circumstances of the attribution being clear for all parts. In other words, there are no facts that determinate the truth of the attribution. In this section, I argue that some of those disagreements involving ascriptions exhibit an action-guiding component, that is, the disagreement reveals that belief ascriptions involve a component that is not descriptive[6]. This argument is anal-

[6]The main argument I present in this section has been developed together with Eduardo Pérez Navarro, Manuel Heras-Escribano and Javier González de Prado (Pérez-Navarro et al., MS)

ogous to Field's (2009) argument against descriptivism in meta-epistemology. According to this argument, appealing to empirical facts cannot solve some disagreements involving knowledge ascriptions. I expand upon Field's ideas in order to draw similar conclusions for desire and belief attribution. As a consequence, attributing mental states cannot be a matter of describing the psychological reality of an agent.

3.1 Disagreement and Descriptivism

According to (Dennett, 1978a, 1987), propositional attitude ascriptions cannot be anchored in internal states of a subject because ascriptions are always subject to a certain kind of 'indeterminacy'. The patterns that support propositional attitude ascriptions "fall short from perfection, as they always must, there will be uninterpretable gaps; it is always possible in principle for rival intentional stance interpretations of those patterns to tie for first place, so that no further fact could settle what the intentional system in question really believed" (Dennett, 1987, 40). As McCulloch (1990) claims, Dennett is "denying that there are deep facts concerning what people really believe or desire" (p. 3).

Dennett (1978a) makes his point with an example. He invites us to consider the case of Sam, an art critic who has promoted the paintings of his son. In principle, there are two possible interpretations of the situation: "a) Sam does not believe the paintings are any good, but out of loyalty and love he does this to help his son, or (b) Sam's love for his son has blinded him to the faults of the paintings, and he actually believes they are good" (Dennett, 1978a, 39). Now, suppose for the sake of the argument that we have a reliable way of determining the cause of the action. Imagine, as Dennett says, that we have the technology to write in Sam's brain a specific judgment. Imagine that we write 'my son's paintings are great' at the moment he is promoting his son's paintings. In fact, we can suppose that this was the occurrent cause of the action (promoting his

son) at this moment. Dennett's point is that, even in this extreme case, we have no deep facts we can appeal to in order to decide whether or not the ascription of this belief is certainly explanatory of the situation. Someone could examine the past and future circumstances of Sam and suspend the interpretation that Sam believes his son's paintings are good. The interpreter could examine Sam's past behavior and realize that he systematically avoided interpreting his son's paintings using the same aesthetics standards that he used for other artists, or that his subsequent behaviors are incoherent with the decision of promoting his son's paintings. These circumstances would give to the interpreter reasons to change his verdict. At the same time, the other interpreter could insist that the accurate ascription is the one that identifies the real cause of the behavior. However, it is dubious that we can decide which belief ascription is right on the basis of all relevant facts. Both interpreters could agree about all the relevant facts and differ in their ascriptions. Furthermore, notice that this disagreement could persist even if the interpreter is Sam himself. Arguably, there are many situations where expressing a belief is enough to ascribe it. However, sincere expressions of beliefs are not always reliable. As Dennett writes:

> This suggests that even if we were to discover a brain-writing system that represented our judgments, the mind reading that could be accomplished by exploiting the discovery would not uncover our beliefs. To return to the case of Sam the art critic, if our neurocryptographer were able to determine that Sam's last judgment on his deathbed was, "My consolation is that I fathered a great artist," we could still hold that the issue between the warring hypotheses was undecided, for this judgment may have been a self-deception (Dennett, 1978a, 49).

Although interpreters usually take the word of the interpretee as a determinant factor, this is not always the case. Occasionally, an interpreter may consider that the interpretee is confabulating, that her judgment is biased, etc.

The case of disagreement has been widely discussed in philosophical areas as meta-epistemology, meta-ethics or semantics of taste predicates (Cohnitz and Marques, 2014; Chrisman, 2007; de Sa, 2015; Field, 2009; MacFarlane, 2014).

In order to appreciate the problem that those types of disagreement present for FP-descriptivism, let me consider the debate in the field of meta-epistemology. The classic descriptivist analysis of sentences of the type 'S knows that P', sometimes called epistemic invariantism, claims that those sentences describe a dyadic relation between a subject and a proposition, and this relation holds independently of other factors or elements. Within this view, such sentences are true or false simpliciter. The problem with this approach lies in the apparent cross-context variability in the truth-values of knowledge attributions. For instance, accepting invariantism would force us to accept skeptical arguments as the following:

P1: I don't know that I'm not a brain in a Vat
P2: if I don't know that I'm not a brain in a Vat then, I don't know that I have hands
C: I don't know that I have hands

If knowledge attributions are true simpliciter and I can substitute the sentence 'that I have hands' for whatever sentence that everybody would agree as first-hand knowledge, then invariantism seems to be forced to accept the skeptical conclusion that we do not know anything.

A way to solve paradoxical conclusions of invariantism is to embrace epistemic contextualism (DeRose, 1992, 1995; Cohen, 1988, 1999). According to epistemic contextualism the same knowledge attribution can express different propositions in different contexts of utterance. One way of modelling this context-dependence is by positing a hidden indexical. Thus, the truth-values and the meaning of a certain claims should be analyzed taking into account that hidden indexical. As applied to knowledge claims or attributions, epistemic contextualism analyzes the claim that 'S knows that p' as having the logical form 'S's belief that p meets epistemic standard e', where the value of 'e' depends on the context in which the claim is uttered (Chrisman, 2007, 226). Thus, both the

proposition expressed and the truth-value of knowledge attribution change depending on the value that the hidden indexical acquires in the different contexts in which the sentence is uttered. This allows one to relativize the proposition according to the epistemic standards, and thus, to restrict skepticism to those contexts where epistemic standards are so demanding because the case that we are brains in a vat is relevant. Changes in truth-value depend on the different conversational contexts in which the claim is uttered.

Epistemic contextualism, as invariantism, is a version of descriptivism. The main difference is that contextualism includes in the analysis certain parameters that make the attribution relative to epistemic standards. However, this inclusion does not alter the descriptive function of the attribution. Now, epistemic contextualism is also problematic due to the nature of the relativization. The problem is made explicit when we treat cases of disagreement concerning knowledge attribution:

> Speaker A: S knows that p
> Speaker B: S doesn't know that p

For instance, imagine two persons disagreeing about whether or not a third person, Lola, knows that the class starts at six. According to the contextualist treatment, the relativization of the truth of the attribution to the epistemic standards of the attributer can be made explicit as follows:

> Speaker A*: S is entitled by epistemic norms eA to her true belief that p
> Speaker B*: S is not entitled by norms eB to her true belief that p

The problem is that, according to contextualism, we can consider that both A and B can be justified in saying what they say by claiming that the contents of A's and B's utterances are different. In other words, an apparent disagreement

between two speakers can be dissolved when the standards are made explicit because A and B would be saying different, compatible things. Once the standards are made explicit, the initial appearance of disagreement dissolves (see Chrisman 2007, 228-230; MacFarlane 2014, 8-13). The problem is that it seems plausible to assume that two speakers attributing knowledge could disagree even though through the conversation they make explicit different epistemic standards.

Disagreement does not always dissolve after making different standards explicit (Field, 2009); rather, disagreement remains, but at the level of supportive attitudes. According to Field, two persons can disagree about ascribing knowledge to someone even if they agree about all the relevant facts involved in the situation. This persistent disagreement demonstrates that knowledge ascriptions are not straightforwardly factual. In other words, knowledge ascriptions cannot be descriptive because there are non-factual components involved in them. Consider the scenario presented above again:

> A: Lola knows that the class starts at six.
> B: Lola doesn't know the class starts at six.

Now, imagine the two speakers have different epistemic standards. B has high-level standard attribution, and he thinks Lola does not know it because she is quite absent-minded. However, A has low-level standards, and he thinks Lola knows it:

> A: Lola is entitled by the norm eA to believe that the class starts at six
> B: Lola, is not entitled by the norm eB to believe that the class starts at six

Field argues that in our everyday life we can find cases like that where the disagreement persists. Notice that A can agree that according to eB Lola doesn't know and B can agree that according to eA, she does. In fact, they can agree

about all the relevant facts, even they could agree about having different standards. However, the discussion could persist. One of the attributers could accept that according to the other's standards there is a different ascription but still disagree about the ascription. The persistence of this kind of disagreement seems to point out that knowledge attribution has a non-factual component, an action-guiding component manifested in the endorsement of the norms. The terms 'action-guiding component' try to capture the idea that the attributer of the attribution is supporting a particular standard that specifies her commitments or authority to defend what is followed from those standards; and also, the commitment to act in ways that are compatible with them (Chapter 5). The reason why the two interpreters could disagree is because they have different attitudes of acceptance towards the norms concerning the attribution. The interpreter could even accept that according to different standards, different ascriptions follow; but still they could disagree about endorsing one standard or another. This disagreement is neither about facts nor about norms. The two speakers can recognize that their interlocutor is supporting a different norm. However, the disagreement does not move into a disagreement about which is the appropriate norm. The disagreement is about whether or not we can attribute to Lola the merit of knowing something, it is a disagreement about the evaluation. Thus, knowledge ascriptions cannot be descriptive.

3.2 Disagreement and Folk Psychology

Now, my point of contention is that Field's argument seems to apply to propositional attitudes broadly considered. Arguably, disagreements concerning belief and desire attribution can be non-factual disagreements in Field's sense. For instance, consider again the example from Dennett. The two speakers disagree about whether or not we can attribute to Sam the belief that his son's paintings

are good. Now, imagine that through the conversation the two speakers make explicit different standards of attribution. For instance, Speaker_1 appeals to a norm of coherence; Sam, she argues, does not usually promote the type of art his son's art paintings represent. Thus, he cannot believe that they are good. On the other hand, imagine Speaker_2 appeals to a norm of credibility. He defends that Sam is a well renowned critic who would not risk his career by promoting bad pieces of art.

What we have here is a disagreement of the type Field presents. Even when both speakers agree about the relevant facts and they recognize that according to different standards different attributions follow, they can disagree because they resist abandoning their supportive attitude to the norm (coherence/credibility). Thus, like knowledge attributions, belief attributions have an action-guiding component supporting standards of attribution. Even if the two speakers make explicit their differences in the norms of attribution, the disagreement does not necessary dissolve. It still makes sense to say: "taking coherence into account, Sam does not believe that the paintings are good; taking credibility into account, Sam does believe the paintings are good, but does he believe it or not". The source of disagreement is not a worldly aspect we describe; rather, it is a supportive attitude component the speaker indicates with the attribution.

Similarly, desire attribution can be read in non-descriptive terms. Consider the following example presented by Hutto:

> Imagine that a friend, F, avows that her reason for going to a particular cinema is that it has started showing art-house films. But suppose you have ample prior evidence that on the whole F despises such films and that she would normally avoid that particular cinema because it is in a bad part of town. Of course, one has to make allowances for changes of mind. But suppose you also discover there is good evidence that F has developed a special but unacknowledged attraction for one of the ticket sellers at the cinema. Moreover, suppose F continues to visit the cinema in question even after it ceases to show art-house films. In such a case, there are strong grounds to doubt F's preferred explanation of her actions. It may be that

> F was knowingly dishonest about her reasons, or perhaps she was self-deceived. (Hutto, 2013, 590).

Again, we have a disagreement between two ascriptions. We can conceive F's behavior as being motivated by the desire of seeing the ticket seller at the cinema or not. The way Hutto describes the case, it seems we have two different standards to attribute the desire to F. A standard of coherence in contrast to the sincerity of F when she avows her reasons. The problem of this case is that Hutto presents it as a case of confabulation and he assumes that F is self-deceived. One may argue that if we know all the relevant facts about the personal history of F, then we can determine whether or not she is confabulating, and therefore, we can determine if F desires to see the ticket seller or not. The problem with this counter-argument is that it assumes that a norm of attribution can always provide the best interpretation. Considering that knowing all the relevant facts concerning the history of F determines which desire motivates the attributee, it assumes that coherence is the best norm to interpret the case. However, one may consider to support other norms, for instance one may appeal to sincerity for supporting F's interpretation. In fact, one interpreter could agree with all the relevant facts of the history of F, agreeing that F is incoherent if she goes to the cinema, but still, supporting that F goes to the cinema because she wants to see art-house films, for instance, by arguing that F changed her mind.

Summing up, the types of disagreements presented in this section have two special features. Firstly, the two interpreters agree about the relevant facts concerning the attribution. Secondly, the disagreements do not necessary dissolve when the standards are made explicit for the interpreters. This points out that propositional attitude ascriptions reveal an action-guiding component, a component that captures the interpreter's supportive attitude to different norms of attribution. Thus, propositional attitude ascriptions cannot have a descriptive nature since they manifest the attributer's supportive attitude. In other

words, FP-descriptivism cannot capture the meaning of propositional attitude attributions.

4 Explanation or Justification?

The aim of this section is to argue that propositional attitude ascriptions do not play the anticipatory/explanatory role that the received view has assumed. Paradigmatically, explaining actions in terms of propositional attitudes is a reactive response to rejections of communicative acts, potential objections or questionings about behavior. Explaining behavior in terms of propositional attitudes is usually restricted to contexts that require evaluating an agent as being responsible for a particular content in order to condemn or justify the course of action[7]. Thus, if ascribing a propositional attitude amounts to evaluating a subject, then FP-Descriptivism is an inadequate model of ascriptions.

Recently, a great number of scholars have defended that propositional attitude ascription is not as pervasive in social situations as the received view has supposed (Andrews, 2012, 2015; Fernández-Castro, 2015b; Gallagher and Hutto, 2008; McGeer, 2007, 2015; Millikan, 2004; Morton, 2002; Strijbos and de Bruin, 2012a,b; Cleave and Gauker, 2010; Zawidzki, 2008, 2013). On the contrary, they support the claim that propositional attitude ascriptions are restricted to contexts where anticipatory capacities fail, that is, when the target's behavior violates the expectations of the attributer. Millikan expresses the idea as follows:

> we expect people to exhibit behavioral patterns similar to those they have shown in the past. Some people usually come to work on foot and on time, others drive or take the metro and often arrive late. Some people always eat lunch at noon, others at other times or irregularly. Some people will talk on and on if you start conversing with them, others are very reticent. Some always stick to their

[7] As I said before, the evaluative component differs depending on the propositional attitude concept in question. However, I will keep talking about responsibility for the sake of simplification

> word, others change their minds frequently. Some always eat eggs for breakfast, others always eat yogurt. We take these patterns into account, betting on their continuation when it is useful or necessary to do so. When we use belief-desire psychology, it is almost always for explanation after the fact, not for prediction. We may explain why John always has yogurt for breakfast by saying he must like it, but if he actually eats yogurt only for his health, it won't matter to our predictions(Millikan, 2004, 22).

Belief and desire ascriptions, Millikan claims, are not used to predict behavior but to explain it when our prediction fails. Before I consider the function of propositional attitude ascriptions, let me consider some arguments against the anticipatory function of mental state attribution.

4.1 Prediction

There are two arguments against the idea that belief and desires ascriptions are necessary for predicting others' behavior. The first argument concerns what Zawidzki (2013) calls the tractability problem (see also Morton, 2002; Zawidzki, 2008, forthcoming). In principle, a particular course of action is compatible with possessing many different mental states. That implies that our mind has to deal with a high degree of underdetermination in order to predict a given behavior:

> Human interactions are too complex to succumb to folk psychological prediction because human decision-making is usually strategic: to predict what other agents will do on the basis of accurate mental state ascriptions, agents would have to take into account what other agents think that they will do, what other agents think that they think others will do, etc. Such an intractable spiral of higher orders of intentionality would inevitably swamp interpreters seeking to predict behavior on the basis of mental state ascriptions. Humans also usually pursue outcomes that are defined in terms of the motives and other mental states of their fellows. This makes preferences inherently unstable. Because the process of making and enacting a decision may reveal or even change others' mental states, preferred outcomes change as one pursues them (Zawidzki, 2008, 145).

Giving this degree of under-determination, attributing a set of mental states to anticipate a course of behavior is an intractable enterprise. A pair belief/desire

is consistent with many different situations and behaviors. Thus, anticipations in terms of beliefs and desires would produce systematic failures of prediction. This problem is more pressing if the holism of the mental is considered. Even if we can make a more or less accurate attribution of desires and beliefs to an agent, there is a high probability that the agent has other mental states (preferences, emotions, etc.) that inhibit the action we anticipate.

The second argument concerns empirical evidence in social psychology. As Andrews (2012) argues, empirical psychology demonstrates that rather than relying on propositional attribution, our predictive capacities seem to be driven by what agents ought to do according to norms concerning situations, stereotypes and social rules. For instance, we categorize people according to social roles or gender and exploit the information regarding this categorization in order to produce expectations (Greenwald et al., 2009; Olivola and Todorov, 2010; Clement and Krueger, 2002). Female infants are expected to be more vulnerable in some situations than male infants, and we tend to associate different roles with each gender (Golombok and Fivush, 1994). Then, we exploit what Kalish and Lawson (2008) called 'deontic relations': information about what a person should do or be like depending on this category. Stereotypes are one of those normative structures that police our interactions. It deserves mention that those social categories are not based on inductively inferred knowledge. We do not treat males and females differently because of differences in behavior; we treat them as we do because we assume they should behave according to gender category. Other sources for anticipation rely on social norms: civic standards, etiquette rules, traffic norms, cultural norms, and so on. These norms facilitate our interactions by enabling people to anticipate what others will do on the basis of what they should do according to them. Maibom (2007) explains this point as follows:

> Consider how people behave in restaurants. What the person who waits on guests does, he does qua waiter; his desire to take somebody's order is a function of him seeing himself as a waiter and is

> quite independent of his personal desires and preferences generally. When, perusing the menu, the other person sees him approaching, she infers that he is coming to take her order, but to do so she need only understand that this is what waiters do with customers and that he is a waiter and she is a customer. What he, personally, desires is irrelevant to the customer's prediction of what he will do and what she ought to do and vice versa.(Maibom, 2007, 568).

Cultural norms regulate our social interactions, so we expect people to behave according to them. This information can be exploited to produce predictions about others' behavior without postulating any mental entity.

Finally, humans anticipate others' behavior throughout the circumstances in which the action is performed Heider (1958). Some of the norms regulating our behavior attend to general standards of rationality. By assuming agents are rational, we anticipate their actions depending on how they should behave in accordance with the circumstances. Traditionally, this capacity has been understood in terms of beliefs and desires (Dennett, 1987). However, as Zawidzki claims:

> Such interpretative competence does not require speculating about concrete, unobservable causes of behavior or appreciating that these causes are full-blown propositional attitudes, that is, states with content represented via individually variable modes of presentation and holistically constrained influence on behavior. It requires only a sensitivity to certain abstract properties of bouts of behavior, namely, that they aim at specific goals and constitute the most rational means to those goals given environmental constraints. (Zawidzki, 2013, 15).

We can generate expectations about others' actions without taking into consideration the inner states of our targets[8].

Summing up, we have reasons to think that belief and desire ascriptions do not play the important role in prediction the received view has supposed. Rather than describing mental states for the sake of prediction, humans seem

[8] A possible objection to this approach would be to defend that those types of normative strategies deployed in social anticipation cannot be acquired without a previous understanding of mental states. However, Zawidzki (2013) (see also Mameli, 2001) has convincingly argued that humans are equipped with certain *mindshaping mechanisms* that facilitate the acquisition of norms and complex patterns of behavior that facilitate social interactions without the necessity of complex mindreading capacities (see Fernández-Castro, 2015a, for a critic).

to anticipate others' actions by assuming they will follow certain rational and social norms. In other words, we exploit social heuristics concerning normative standards about how social creatures ought to behave.

4.2 Explanation as a form Justification

Now, the question is whether or not propositional attitude ascriptions play a pervasive role in explanation. Notice that according to FP-descriptivism, failures of anticipation must be taken as failures in the process of interpretation, that is, failures in the postulates of ascriptions or the information of the context. Given that, one may expect the attributers to react by revising the process, for instance, replacing the mental state postulated previously by a more accurate one. When people behave contravening our expectations, we must have failed somehow in our descriptions of the states or inferential processes. The metaphor of the scientist is especially illuminating in this case. As a scientist who fails to anticipate the results of an experiment, a folk psychologist can fail in her ascriptions, in considering the variables involved (contextual information) or in the inferential process.

The problem with this picture is that humans do not always react to failures of anticipation in this way. Our reactive responses do not only include explanations, we respond to failures of anticipation by deploying other restorative strategies: sanctioning the behavior, asking for reasons, or excusing the behavior of the target. This point has been strongly emphasized by several authors (Andrews, 2015; McGeer, 2007, 2015; Zawidzki, 2013). For instance, McGeer says:

> In my view, what is most noteworthy in these cases is the fact that folk psychologists have, as part of their overall competence, myriad techniques for identifying, excusing, blaming, accepting responsibility, apologizing and otherwise restoring confidence in the efficacy of the normative structure that govern the behavior of individuals who ought to be explicable and predictable using the techniques of folk

> psychology, even though sometimes they are not, in other words, folk psychologist treat lapses of rationality, not just as "surd spots" in an explanatory/predictive theory, but as reasons to take some kind of remedial or restorative action (McGeer, 2007, 142).

Rather than exhibiting failures of the interpreter to describe the accurate inner causes of the behavior, our profiles of responses show that we take those failures as anomalies in the interpreters' capacity to deal with the social situation. At least in some cases, it is the target who is failing to perform what she ought to do. Explaining the behavior is not the unique response we exercise when our expectations fails. A terrible reminder of those regulative practices is the type of responses that society deploys to make others conform to gender or ethnic norms (micro-inequities). For instance, as the Chilean rapper Ana Tijoux comments in an interview for feminist magazine Pikara: "It never ceases to amaze me that they still ask me who takes care of my children when I go on tour. This is a question they do not make to men who are fathers" (Pikara Magazine, 2015). Stereotypes generate certain expectations as normalized behaviors, ways of thinking and feeling and in general, a set of oughts to which the person must conform due to being categorized in the way she is[9].

Of course providing explanations is an important part of those restorative responses, but they are bound to the elucidation of counter-normative behavior as well. As Andrews puts it:

> Given an understanding of norms in a society, and the ability to recognize and sanction violations, there developed a need to understand

[9] These approaches to folk psychology usually focus on how our regulative practices facilitate the maintenance of norms and rules that governs our social interactions. However, some of those practices are also important tools for social change. For instance, providing reasons does not only help us to justify or excuse counter-normative behaviors, but also, to make our interpreter recognize that was correct to break the norm or substitute it for another one. The emphasis on social structures and anti-individualistic mechanisms as the sources of social dynamics are not only explanatory of social behavior (Haslanger, 2015); but also, it helps us to recognize the importance of transgression and resistance for subverting unjust social situations. A hopeful conclusion of the regulative approaches is that, as radical movements have powerfully revealed, acts of transgressions such as the resistance of Rose Park to stand up on a bus in Montgomery can change social perception. Social transactions, as other normative practices, are subjected to, and often must, change.

> actions that violated the norms. Explanations for norm-violating behavior that didn't cite a person's reasons either led to excluding the individual (e.g., "He fed because he is crazy, so let's stop sharing meat with him"), or they failed to satisfy those who demand an explanation. This need to have a satisfactory reason for the behavior of one's companions is what drives the need to develop another sort of explanation, namely reason explanations. There is a significant benefit to being able to explain behavior that violates norms, because explanations of the right sort can also serve to justify behavior.(Andrews, 2009, 445).

In contexts where social understanding is governed by norms, it makes sense to have the possibility of justifying behavior and making it understandable in order to avoid public sanctions. Reason explanations are also a possible *reactive response* when our behavior is perceived as anomalous or when we are encouraged to exculpate ourselves.

When someone questions our actions or speech acts, when someone asks us for our reasons to act or notices our incongruences, we provide different reasons to exculpate our actions. In other occasions, we attempt to exculpate others for something they did, or we try to find a reason to condemn their behavior because we consider it immoral or inappropriate. In general, providing reasons is a tool for social cover when facing the possibility that our actions could be open to sanction. Propositional attitude ascriptions are reasons used as *reactive responses*[10] as well. For instance, one may need to appeal to propositional attitudes when someone questions our actions: 'Why are you not dressed up yet?' 'I thought it was earlier' or 'Why did you get up so early?' 'I want to go running before going to my class'. Other situations that demand the use of propositional attitude ascriptions are those contexts where someone notices our errors and we need to excuse the mistake: 'I came late because I thought the film started at six'; 'He believed that he could make a better job'. In other situations, we want to indicate our degree or lack of support concerning an assertion to anticipate

[10] Other paradigmatic uses of normative concepts as reactive responses are truth attributions (see Ramsey 1927/1991, 12 and Frápolli 2013, 79) or ascriptions of irrationality (Hayward, 2016).

possible negative reactions: 'Propaganda of the deed, I believe, is a fair political action'; 'Socialism, Chávez believed, is the best solution for Latin-America'. Other times, we want to do the opposite, for instance, reacting to a possible disagreement with a challenging tone (I believe in evolution). In any case, all these examples share the basic function of reacting to possible indications of violations of social norms, that is, the possibility that a certain course of action is contravening a norm. In fact, in a recent empirical study, Korman and Malle (2016) have shown that people offer many more reasons in terms of mental states when they face puzzling actions in contrast to ordinary actions. In these experiments, a group of participants were presented with situations where behaviors were "puzzling with respect to social perceivers' prior knowledge and expectancies about behavior in general" (p. 3), in contrast to another group presented with ordinary behaviors. Although both groups provided a similar number of explanations, the group presented with puzzling behavior tended to provide more reason explanations (in contrast to trait or causal explanations) and mental state explanations than the other. This seems to provide support for the idea that reason explanation in general, and mental state attribution in particular, are reactive responses to anomalous behavior, those behaviors that contravene normative standards governing social situations[11].

Another way to notice that propositional attitudes are reasons governed by normative structures that help us to normalize behavior is that they are not always acceptable in social circumstances. Reasons are not always permissible given social standards (Tanney, 2013). Suppose a firefighter is ready to run into a burning building, but she flies out the building. Now, imagine we ask the firewoman why she fled. She answers she wanted something to eat, and so left the building. Consider how awkward we would find this answer. As Tanney

[11] Of course, this does not mean that we only use propositional attitude ascriptions in those contexts. We evaluate others' person with other purposes apart from exculpating or condemning them. I discuss this issue in the following chapter.

says:

> we would reject this as an explanation on the grounds that it does not make sense. Supposes she says that it makes perfect sense to her why she would drop her everything—even put lives at risk—because she wanted something to eat. We just do not understand what it is like for her when she wants something. Indeed, she is right: we do not understand (Tanney, 2013, 143-144).

The reason why such an explanation would not make sense to us is because according to our standards, saving persons from fire is more valuable than eating. Propositional attitude ascriptions are governed by standards of normalization.

In order to justify or condemn a pattern of behavior we need to consider the subject responsible for his actions. In this sense, those uses of ascriptions have an evaluative function. In those contexts, when we ascribe a propositional attitude we are burdening the subject with the responsibility or credit derived from undertaking the content. Those uses of propositional attitude ascriptions can only play the role of justifying or condemning a particular behavior because they have the function of assessing the subject as someone who is committed to what is followed from the content. FP-descriptivism presupposes that mental verbs describe the psychological states of the attributees. However, considering attributions as descriptions of psychological states does not capture the action-guiding component toward the reason. Notice that this does not mean that our ascriptions do not have descriptive uses (see Chapter 5). However, my point of contention is that ascriptions of propositional attitude in folk psychological situation, like those mentioned above, are paradigmatically used with evaluative purposes. In those circumstances, using a reason that includes a propositional attitude involves a specification of how to evaluate the agent. In particular, her position toward the reason and specification about the behavioral responses of the agent. For instance, when someone says 'the firefighter came into the building because he thought there was someone inside' we are making the firefighter

responsible of the content (there was someone inside) in order to justify the action. As I argued in chapter 3, the type of action-guiding information specified by the ascriptions seems to differ from the type of information provided by descriptive-states or expressions. In general, the function of these propositional attitudes is required in cases where we need to explore responsibilities, degrees of approval, convictions and, in general, when we take an evaluative stance to the behaviour of our interpretee. This evaluative stance does not seem to be captured by folk psychological descriptivism.

4.3 Reasons, Facts and Evaluation

A possible objection to this view is to recognize that propositional attitude ascriptions may involve evaluations, but arguing there are other types of reason explanations with same evaluative component which do not involve propositional attitudes. Thus, what distinguishes ascriptions is not their evaluative character but their reference to inner psychological states of the subject. In order to answer the objection, let me unpack the argument. An important part of our explanations does not require to make reference to propositional attitudes. These reasons are *factive* (Strijbos and de Bruin, 2012a,b, 2013a), that is, they are reasons of the type 'S performs the action because of P' where P states facts and values enabling or entitling the behavior. For instance, we say that Mary ran away because the building was on fire or that Anthony went out to take a real espresso. Although these reasons do not appeal to mental states, they introduce the evaluative component specified before, we evaluate a person by introducing a fact that plays a role in the justification of the behavior. Those facts entitle or support the action on the basis of certain standards (Bruner, 1990; Hutto, 2004, 2008a; Perner and Roessler, 2012; Strijbos and de Bruin, 2012a,b). In section 4.1, I argued that our anticipatory capacities are generated by different types of social norms and structures that help us to anticipate what others will do on the basis of what they ought to do. Making sense of others

does not require exercising any mentalizing capacity by us. Alternatively, "The work is done and carried by the world, embedded in the norms and routines that structure such interactions" (McGeer, 2001, 119). In a similar vein, we can argue that our factive reasons are maintained by the same normative structures. We do not need to describe the inner mental states of the subjects, rather than describe the contextual details that make the action permissible given those normative structures. Hutto explains the idea as follows:

> [Reasons] function as 'normalising' explanations, allowing us to cope with 'unusual' or 'eccentric' actions either by helping us to see them as familiar or by making them so. This is achieved either by supplying missing details that reveal an action to be in the fold of the ordinary already—despite appearances—or by fleshing out a larger context such that we come to find it acceptable. This use usually entails that we extend the range of what we think of as falling within the scope of the 'normal'. But it goes without saying that this sort of 'negotiation' requires a prior fluency with 'the normal'. Hutto (2004, 560).

Elucidating and normalizing behaviors implies attending to non-obvious events and circumstances that make the behavior normal. Reasons fill the missing details or help to appreciate the circumstances surrounding the action in order to explain it. As Strijbos and de Bruin put it:

> "the basis for our understanding of others lies outside the mind of particular agents, in the context of a shared practical world. When we try to make sense of the actions of others, one of the first and most important tasks is to figure out what it is in this shared habitat they are responding to. And the first step of the agent who is asked to explain his action in terms of reasons is precisely to provide this fact.(Strijbos and de Bruin, 2013a, 163).

These explanations involve pointing out and tracking the relevant events that normalize the target's behavior, without explicit reference to her inner states.

For our purpose, the importance of reason explanations lies in the fact that they introduce the same evaluative component discussed above without explicit use of mental states. The genuine function of reason explanation is not describing a fact, but situating the agent as someone from whom it makes sense

to expect the explained behavior given the situation where he is situated. Of course, there is a fact involved (the building being on fire); nonetheless, the reason functions to make the agent responsible of the behavior given his acknowledge of the fact that normalizes the behavior given the social norms that rule what to do given the situation. Having said that, one may exploit the disanalogy between factive reasons and reasons in terms of propositional attitude in order to argue that the difference lies in the descriptions of psychological facts of the propositional attitude ascriptions. In other words, while the evaluative function is shared by all types of reasons, the distinction between factive and mental reasons resides in the description of the inner psychological reality of the subject.

Traditionally, factive reasons are considered mental in nature. The received view recognized that explanation in factive terms are quite normal in our social interactions. However, it has assumed that the description of mental states is implicit in the process. Fodor favors this interpretation when he says: "subsumption under platitudes is not the typical form of commonsense psychological explanation. Rather, when such explanations are made explicit, they are frequently seen to exhibit the 'deductive structure' that is so characteristic of explanation in real science" (Fodor, 1987, 7). Fodor seems to think that factive reason explanations hide deductive inferential processes involving propositional attitude concepts. In other words, factive reason explanations mirror formal inferences and include propositional attitude ascriptions. In this view, inferences such as 'the woman ran out of the building because it was on fire' derive from explicit formal inferences such as 'the woman thought the building was on fire, she didn't want to burn, she believed running would make her staying safe, thus, she ran out of the building'.

From my view, the aforementioned factive reasons are evaluative. In this

sense, I believe there is a grain of truth in considering factive explanation is a question of mentalizing. What Fodor does not realize is that providing a folk psychological explanation is to evaluate a person as acknowledging the situation and what is behaviorally connected with it. In a sense, we are giving a special status to this person, we are considering what she must do as a social creature with certain duties and commitments. But this does not require appealing to her mental profile or describing her internal states. Human anticipatory and explanatory capacities do not require describing propositional attitudes or psychological states of the target. This normalizing strategy does not require describing mental states, rather than evaluating the subject as someone responsible for the action and who acknowledges the normative connection between the situation and the action she performed.

Now, the question is what makes propositional attitude different from factive reasons, i.e., what is special about ascribing propositional attitudes and their function to navigate social situations. In order to see the contrast, consider the following examples presented by Malle (2004):

(14) Why did he refuse dessert? He's been gaining weight.

(14') Why did he refuse dessert? He thinks that he has been gaining weight.

(15) Why is she taking the car to work? Because she is late.

(15') Why is she taking the car to work? Because she believes that she is late.

(16) Why did Ben call Anne? They would make up again.

(16') Why did Ben call Anne? He hoped they would make up again.

(17) Why did she go to the coffee shop? To have an espresso.

(17') Why did she go to the coffee shop? She wanted to have an espresso.

When contrasting (14-17) with (14'-17'), the difference between the explanations seems to lie in how specific they are about the commitments undertaken

by the evaluation. For instance, by uttering (14), the attributer assesses the target as being responsible of gaining weight, which counts as a reason to support the behavior. On the contrary, by uttering (14'), the attributer is being more specific, he is indicating that is only the attributee who is responsible of those commitments. This allows the attributer to remain neutral with respect to the acceptance of the fact that the target is gaining weight. This is the reason why it would be weird to say 'He's been gaining weight but I am not sure'. But it would make sense to say 'He thinks he is gaining weight but I am not sure'. Factive reasons are a default form of knowledge attribution (Gordon, 2000; Strijbos and de Bruin, 2013a), they presuppose the subject share the same knowledge of facts about the world the attributee is considering. A similar claim may be supported concerning the other examples. Propositional attitude verbs such as 'guess', 'suppose', 'know', 'want' manifest the attributer's evaluation concerning the attributee's position regarding the content. That is, different degrees of approval, disapproval, conviction or other attitudes towards the reason accounting for the behavior of the agent. The reason why we introduce mental verbs instead of using factive reasons is to be more specific concerning the commitments. This is not always necessary, in fact, introducing mental verbs in some occasions where it is not necessary can sound strange or redundant (see chapter 5). But the point is that what makes reasons in terms of propositional attitudes distinctive is not that they describe the inner psychological reality of the subject, but the specifications concerning the evaluation.

In a nutshell, regularly, reason explanations are tools to justify or condemn particular patterns of behavior when they contravene a norm or deviate from a normal pattern. They can provide this function in social situations because they are evaluations, that is, they indicate some merit, credit or responsibility of the subject toward a content. Propositional attitude verbs function as tools to

give specification about the evaluation. They introduce information about how the subject must guide her behavior according to the evaluation. This action-guiding information is substantially different from the descriptive information about the world that descriptive expressions introduce. Thus, FP-descriptivism cannot account for what we do in our social circumstances by deploying propositional attitude ascriptions.

5 Conclusions

In this chapter, I have presented three arguments against FP-descriptivism. According to the first argument, there are at least two uses of propositional attitude ascriptions (parenthetical and communicative uses) that seem incompatible with a descriptivist reading. Secondly, I argued, following Field's (2009) analysis of normative disagreements concerning knowledge, that disagreement concerning belief and desire attributions exhibit an evaluative component that is not captured by the traditional descriptivist analysis. Finally, I have argued that propositional attitude ascriptions are bound to contexts of justification/condemnation where the speaker evaluates the attributee concerning his responsibility, credit or merit toward a content. Again, the conclusion is that FP-descriptivism cannot account for this expressive component.

These three arguments seem to me compelling enough to motivate an alternative to FP-descriptivism. The following chapter is devoted to present such an alternative: The evaluative view. After that, I explore different consequences of this alternative for the relation between language and mentalizing. In particular, I will argue that acquiring the capacity to evaluate others requires to engage in joint activities conversationally mediated during development (Evaluative Conversational Hypothesis).

Chapter 5

The Evaluative View

1 Introduction

In the previous chapter, I have argued that paradigmatic cases of propositional attitude ascriptions in folk psychological situations are non-descriptive: these uses of propositional attitude ascriptions do not describe the private psychological machinery of the attributee. Furthermore, in contrast to the received view, I have maintained that paradigmatic contexts of propositional ascriptions include those of justification/condemnation of actions or speech acts, exculpation and disapproval of actions. The presence of propositional attitude ascriptions in those situations reveals an evaluative component that cannot be captured by folk psychological descriptivism.

This chapter aims to provide a minimal characterization of what taking propositional attitude ascriptions as evaluations involves. My purpose is not to give a sophisticated account of the semantic meaning of mental predicates. Rather, I put forward a minimal characterization that allows for drawing the connection between the special evaluative features of propositional attitudes and

the empirical hypothesis I will offer in the following chapter. After characterizing the evaluative view, I will present different motivations and independent arguments in its favor.

2 Non-descriptivism and Assessment

As I explained in chapter 3, descriptivism is a widely discussed general semantic conception. Descriptivism is behind what Austin (1962) called the 'descriptive fallacy' (Ryle, 1949, 56,115; Ryle, 1979, 89; Belnap, 1990, 1). The denial of descriptivism is a well-known strategy in different areas of the philosophy of language in order to approach different expressions and discourses. In this respect, we can find non-descriptivist approaches concerning ethical claims (Ayer, 1936), logical expressions (Wittgenstein, 1922), avowal (Bar-On, 2004), epistemic claims (Chrisman, 2007), semantic concepts (Gibbard, 2012; Lance and O'Leary-Hawthorne, 1997) or rationality ascriptions (Gibbard, 1990). I have put forward different arguments concerning non-descriptivism about propositional attitude ascriptions. These arguments, I believe, are compelling enough to undermine the basic assumptions behind the theories concerning the relation between language and folk psychology I presented in chapter 2. However, one may deny that a particular part of the discourse is descriptive, but consider different approaches about how we must characterize the function that this discourse plays in our social interactions. This move is particularly salient in the different non-descriptivist approaches we can find in meta-ethics. Although such approaches share the basic refusal of the idea that moral discourse denotes moral properties, they differ in the treatment of the function that this discourse plays in our conversations. For instance, Stevenson (1944) argues that ethical claims are signs whose meaning can be characterized as dispositions that relate the sign to a range or attitudes or emotions. Thus, moral claims as

'apartheid is wrong' have the function of expressing a person's negative emotion toward apartheid. On the contrary, Hare's(1952) prescriptivism considers that the function of moral discourse is to express different prescriptions. A prescription entails an imperative, and therefore, assenting to a moral claim is to accept a prescription of an action.

In a similar vein, one may consider different approaches to non-descriptivism concerning propositional attitude ascriptions. In the previous chapter, I have argued that paradigmatic uses of propositional attitude ascriptions are evaluative, that is, they are used to ascribe different degrees of significance, merit, credibility or responsibility to a particular subject. Again, we can find a distinction between evaluative and descriptivist vocabulary in the field of meta-ethics. Expressivism has been characterized as the idea that moral judgments serve to express certain evaluative attitudes toward an object, i.e., the expression 'torturing is wrong' expresses a negative evaluation of torturing. Depending on the expressivist view, one may account for this evaluation in different ways (see Chrisman, 2010). Be that as it may, the aim of this chapter is not to provide a model to capture this semantic aspect of propositional attitude ascriptions. Although one may recognize my position as a sort of expressivism, I am not committed to any particular theory of how the evaluative component of these uses of propositional attitude ascriptions should be modeled. On the contrary, I will introduce a minimal characterization of what is to evaluate a person in the context of social interactions –in particular, what we do when we ascribe a propositional attitude to a subject. This characterization was partially envisaged in the previous chapter. In the next section, I will elaborate upon some of the ideas presented in chapter 4 to give this minimal characterization.

3 The Evaluative View

One of my argumentative strategies against folk psychological descriptivism was related to a particular way of understanding folk psychological practice. In particular, I aligned myself with certain approaches to social cognition (Andrews, 2012, 2015; McGeer, 2007, 2015; Zawidzki, 2008, 2013) that take our social skills to be are somehow distributed in certain normative structures and routines that facilitate our interpretive capacities. According to them, humans are able to coordinate without engaging in complex reasoning strategies because the work is offloaded into those routines and structures which regulate human actions and make them more transparent to one another. Furthermore, our social interactions require tackling situations where agents contravene those norms or routines or when the interpreter perceives the action as anomalous. The way social creatures handle these situations is by displaying different regulative responses that bring the agents back to the normalized actions or make the attributer realize that the action was somehow within the limits of what is dictated by the social norms and routines. These regulative responses include asking for reasons, blaming, and sanctioning.

Propositional attitude ascriptions, along with factive reasons, can be framed into these regulative responses. Ascribing beliefs and desires is a regulative response to counter-normative behaviors, but also to other ways of altering the normative structures of social interactions, such as potential rejections, anticipating possible sanctions, exculpating previous violations, etc. Furthermore, there are other possible purposes for ascribing beliefs and desires, for instance, reducing cognitive dissonance or gaining greater control over a coordinated action (Andrews, 2012, 153). For instance, one may ascribe to his friend the belief that politically incorrect humor is funny because he does not want to admit his friend is a racist. Or imagine the case of a basketball player who ascribes to her

partner the believe that the opponents are bad at defending the pick and roll game in order to exploit this advantage. Ascriptions do not necessary involve to violate a normalized pattern of action or contravene a norm. However, this does not erase their evaluative character, that is, the proposal of someone as a person from whom you can expect certain actions. When we ascribe a belief or a desire to a particular agent we are giving credit, responsibility, merit or significance to her in a particular way. Consider how the basketball player could react if her partner did not start playing the pick and roll. She will be licensed to ask for explanations or even to verbally sanction her partner. These types of restorative responses could only make sense if mental ascriptions were a type of assessment. Comparably, other types of evaluations exhibit the same kind of features. For instance, considering someone a good sommelier implies expecting certain oughts or actions (knowing how to discriminate type of wines, recommending different designations of origin depending on your food), certain skills (olfactory and gustatory capacities); but also, certain responsibilities that allow others to press her when she does not do what is expected. As a result, treating someone as minded, like treating her as a good sommelier, means situating this person in a position such that we can expect from her a set of actions, speech acts, and, in general, a set of behavioral patterns regulated by normative structures.

As I said in chapter 3, we can account for the distinction between the descriptive and evaluative character of some expressions or uses in virtue of the characterization of different types of information. The instances of propositional attitude ascriptions that concern us carry a special type of information; this information specifies how to regulate action rather than representing the world. Following Charlow (2014), we can say that informational content encodes a locational perspective, that is, it helps the subject to locate himself in a logical

space of possibilities by providing information about the environment. On the contrary, action-guiding information is motivational information that helps the subjects to guide their behavior. When ascribing a propositional attitude to a subject, we are providing information about how the subject must behave, not about his or her internal states. Situating someone as believing or desiring is not saying that she is in a particular mental state. As Gibbard (2003) puts it: "When I say he "expresses" a belief, I don't mean he has that belief. To express a state of mind, as I use the term, is to purport to have it, whether or not one does" (p. 77). In fact, as several of the arguments I discussed in the previous chapter have shown, one may avow or ascribe a mental state to oneself without being in it. When an agent says 'I believe P' or 'I want Q', she is entitling others to treat her as someone who is committed to particular patterns of actions or speech acts. However, she is not necessarily saying she is in a particular psychological state. Furthermore, situating a person in the way characterized above cannot be a description of the behavioral inclinations of the subject. When ascribing a propositional attitude, we are making this person responsible for what she ought to do given her situation. That is, when ascribing to Sara the belief that Prometheus is a terrible movie, we assign to her certain obligations, duties and entitlements according to what is normatively connected to considering Prometheus as a terrible movie.

One may worry that considering propositional attitude ascriptions as evaluations could burden our ascriptions and reasons with a moralist or quasi-legal attitude that our social interactions do not always reflect. Part of the motivation to consider ascriptions as evaluations was that they seem to be triggered when we judge an action as anomalous or transgressive. We need to evaluate the subject's degrees of responsibility or merit to situate her in connection with the action. However, one may find situations where we want to explain certain

courses of action that did not need justification/condemnation or did not violate our expectations or norms. For instance, when we are explaining to our friends what we did for vacations or when we just give our reasons for a particular action to keep our partners informed. Our explanations do not necessary involve justifications or condemnations and, thus, it is not surprising that we can find ascriptions and reasons which explain a course of action without necessarily carrying on an assessment.

However, my point here is that these cases, even if frequent, deviate from paradigmatic examples which are evaluative. Reasons and ascriptions are standard tools for evaluating subjects' responsibilities in connection with their actions. And that is the reason why they are normally used in context of providing rationale for actions. One way to recognize this point is by considering how difficult people find to explain actions they strongly condemn (Andrews, 2004, 2012); in particular, when these actions are accounted for in terms of reasons or ascriptions. For instance, when someone offers an explanation for an immoral behavior, e.g., Daesh killed hundreds because of Western countries' foreign policy in Muslim territories; people tend to interpret the explainer to be a supporter or defender of the immoral action. Andrews (2012, 155) provides a good example of this tendency manifested in the condemnatory responses to the intention of the company Alliance Atlantis to produce a TV series about Hitler's early career. Many people worried that the series would be interpreted as a justification of the genocide.

Another way to notice that ascriptions are default tools for assessing subjects in justificatory/condemnatory contexts is reflected in the phenomenon known as *Knobe effect* (Knobe, 2003). Joshua Knobe discovered that people's intentional explanation is influenced by the perception of side effects as good or bad. In the classical studies, two groups were presented with two similar stories whose side

effect differs in moral valence[1]. People who were presented with the bad side effect tended to judge that it was brought intentionally, while the people presented with the good side effect tended not to attribute the intention. Although the classical example involves the attribution of intentions, the phenomenon has been reproduced with other mental states such as desires (Tannenbaum et al., 2007), knowledge (Beebe and Buckwalter, 2010) and belief (Beebe, 2013). It is beyond the purposes of this dissertation to address the explanations of this phenomenon, which has produced a significant amount of literature (see Feltz, 2007for a review) but let me briefly say how it relates to the evaluative view.

The reason why people tend to attribute intentions or beliefs to the person bringing the bad side effect is that the condemnatory attitude toward the action requires considering the person as being responsible for the action. From the view I canvass here, asking for someone's intention is asking for the subject's significance or responsibility concerning a particular action. In the situations presented by the Knobe effect, the positive side effect does not trigger the attributers' ascription because the situation is described in a way that the attributee is avowing his motivation for the action (which is the profit). However, the negative side effect triggers the attribution of intention because the attributee is aware of the immoral effects of his action and thus, he is responsible for it. In other words, we tend to attribute mental states when a behaviour has violated a particular norm (see Uttich and Lombrozo, 2010, for a similar point). Notice also that one may expect subjects to burden the attributee with more re-

[1]The two stories were presented as follows: Harm: The vice-president of a company went to the chairman of the board and said, "We are thinking of starting a new program. It will help us to increase profits, but it will also harm the environment. The chairman answered, "I don't care at all about harming the environment, I just want to make as much profit as I can. Let start the new program". They started the new program. Sure enough the environment was harmed. Help: The vice-president of the company went to the chairman of the board and said, "We are thinking of starting a new program. It will help us to increase profits, and it will also help the environment." The chairman of the board answered, "I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the new program." They started the new program. Sure enough, the environment was helped(Knobe, 2003, 191)

sponsibility due to their condemnatory attitude. In this line, when subjects are allowed to choose between knowledge and belief ascriptions they strongly prefer to attribute knowledge in order to reinforce the responsibility of the attributee with the immoral action (Beebe, 2013).

An interesting way to test the evaluative view would be to check what would happen if the attributers were asked to justify an action in spite of a bad side effect. For instance, forcing the subjects to exculpate immoral behaviors, instead of merely explaining them. The evaluative view would predict that the attributer will use more exculpatory mental state ascriptions in order to put the attributee under a positive light. Although such empirical research has not been implemented, work by Malle et al. (2007) has shed light on this issue. Malle and his collaborators carried out an intensive research concerning intentional explanation and provided empirical evidence demonstrating different asymmetries in how people interpret behavior depending on whether they are actors or observers of the action. In the context of this research, they found that, in general, actors produce many more reason and mental state explanations than observers in comparison with causal history or trait explanations[2]. However, this asymmetry could be reverted when the subjects were explicitly instructed to portray the actor under a positive light. In other words, subjects provide more reason explanations to support others' action when they have to justify it.

In light of this, we are now in a position to make explicit a minimal characterization concerning the function of the relevant uses of propositional attitude ascriptions. I call this characterization the evaluative view:

> **The Evaluative View**: Paradigmatic uses of propositional attitude ascriptions in folk psychological situations do not serve to describe. They serve to evaluate a person as having different degrees of responsibility, merit, or demerit towards a particu-

[2]This is also expectable from the evaluative view. After all, it is natural to assume that we tend to justify our own actions rather than the actions of others

> lar content. This is translated into normalizing someone's past behavior according to certain standards that follow from the content of the propositional attitude, and into situating that person as someone from whom you may expect to behave in ways compatible with such a normalized pattern

Paradigmatically, mental state verbs introduce assessments. In the previous chapters, I argued that factive reasons are also evaluative. However, the difference resides in the specificity of commitments and entitlements that mental states introduce. In order to see the contrast, consider the case of desires. Attribution of desires specify the individual-relativity of the practical reason endorsed. In other words, desire attributions individualize the agent's reason for action. When a speaker says 'Sara went to the pub because she wanted a beer' (in contrast with Sara went to the pub to have a beer), she is highlighting that the practical importance of having a beer or the motivation for having a beer is a practical reason for Sara but not (necessarily) for the speaker or anyone else. Strijbos and de Bruin explain the contrast as follows:

> By ascribing a particular desire or emotion instead of describing the relevant situation as undesirable or having a certain import, the interpreter underscores the merely intrapersonal validity of the practical inference endorsed. He thereby in effect individualizes or particularizes the agent's reason: it is a reason the agent responds to, but he himself (or anyone else in general) need not, under similar circumstances. (Strijbos and de Bruin, 2012b, 155).

In other words, by using the concept of desire, we are emphasizing the individual character of the reason. A similar contrast can be found in the case of beliefs, where third personal ascriptions serve to indicate the responsibility of the subject toward the reason.

Now, the different mental concepts we apply in our social situation reveal different degrees or fine-grained specifications of our evaluations. That is, different mental concepts specify different types of action-guiding information. Consider the following examples:

(1) Sara hoped to have a beer

(1') Sara expected to have a beer

Like desire attribution, (1) and (1') specify the particularized character of the reason. However, unlike desire attribution, (1) and (1') are specifying the degree of commitment that Sara has about the possibility of fulfilling the desire, being this possibility more remote or distant in (1) than (1'). Similarly, when saying 'Sara craves for a beer', in contrast to 'Sara wants a beer', there are different types of expectations concerning Sara's action. We can observe this in the higher effort appreciated when saying 'Sara refrained herself from having a beer in front of her father even though she craved for one'. Again, hope, expect or crave pin down different evaluations of Sara which correspond to different pieces of action-guiding information, i.e., behavioral patterns which are expectable and demandable from her.

In a nutshell, attributing propositional attitudes in folk psychological explanations is to assess a person, to assign her a value that burden this person with credit, merit, responsibility, or significance. Assigning a value gives specifications about how this person must behave; proposing her as someone we can expect to regulate her behavior according to particular normative structures that prescribe what to do in each situation. In this sense, evaluating a person as hoping, desiring or craving is providing information which does not represent any worldly aspect but informs about how this person guides her behavior, i.e., different mental concepts depict different action-guiding information.

There is a worry that may arise fairly quickly, the possibility that we can use propositional attitude ascriptions as descriptions. One may, for instance, call our attention upon cases where propositional attitude explanatory ascriptions do not demand to specify a particular evaluative objective. A possible example would require attributing propositional attitudes while remaining neutral about the action or the situation, for instance, engaging in fictional situations as those

presented by movies, novels or theatre plays. Imagine someone who is watching the TV show Game of Thrones. She sees a furious Tyrion Lannister looking back to the audience of his trial and lets out his rage: "I wish I was the monster you think I am! I wish I had enough poison for the whole pack of you! I would gladly give my life to watch you all swallow it". Now, the beholder knows exactly what is going on there because she has been tracking the narrative of the show:

> *Tyrion's father and sister accuse him of a crime he did not commit. The woman he is in love with is betraying him and she is giving false testimony against him. The nobles of King's Landing, whom Tyrion saved during a city siege, attend his trial to see how he is proven guilty.*

In principle, there is nothing in these events the beholder is tracking that requires her to ascribe a mental state. She just can make sense of the situation because the narrative is coherent with the type of social and rational norms we are schooled to follow. Now, a certain situation that demands explanation could arise. Imagine someone raising the questions: 'why is this guy so angry?' or, 'why is he saying all these things to these people?'. Now, the attributer needs to give some significance or credit to Tyrion, situating him as someone who we can burden with all these reactions. We need to provide factive or mental reasons about his behaviour (he is so hurt because he thought this woman loved him but she is betraying him, or he is furious because his family wants to condemn him). Whatever these reasons are, there is no way of making sense of Tyrion's behaviour without burdening him with responsibilities or commitments derived from the situation. That is the reason why introducing propositional attitude ascriptions when the situation does not demand them sounds awkward (Goldie, 2007, 104). For instance, when we make explicit all the mental states involved when evaluations are not relevant: *Tyrion is mad because he knows that his father and sister accuse him of a crime he knows that he did not commit, and of course, he does not want to be accused for something he knows he did not do.* This does not mean that we cannot use our propositional attitude ascriptions in

descriptive terms sometimes. For instance, when we ask about someone's beliefs about a topic our of mere curiosity or because we want to gain knowledge about the person; when we make generalization judgements like 'Vikings believed that warriors go to Valhalla after dying', 'Pacific cultures believe we cannot speculate about others' minds'; When using ascriptions embedded in larger constructions "according to psychologists, people with Cotard's syndrome believe that they are dead"; or when we talk about our mental inclinations like in 'When I come into the living room, I always think that my mother should have been a interior designer'. One may also characterize as descriptive those uses where we guess or speculate about what the other is thinking when we ask 'do you know what I am thinking' or 'what is le penseur doing'. Admittedly, those and other uses could be genuinely descriptive. However, they seem to derivate from genuine evaluative uses. Ayer (1936) makes this point concerning moral terms as follows[3]:

> (...) I think that what seems to be an ethical judgement is very often a factual classification of an action as belonging to some class of actions by which a certain moral attitude on the part of the speaker is habitually aroused. (. . .) Now in these cases the form of words by which the factual statement is expressed is the same as that which would be used to express a normative statement; and this may to some extent explain why statements which are recognized to be normative are nevertheless often thought to be factual. (Ayer, 1936, 21).

Although folk psychological expressions are paradigmatic evaluative tools, one can expect to use them with other purposes, for instance, describing or classifying. However, they are not the type of ascriptions humans display in genuine folk psychological explanations. Folk psychological explanations are demanded in situations where others' actions become relevant for our purposes, goals or activities, thus, we need to use ascriptions and reasons to charge this person with responsibility, merit and credit that helps us to situate her in a sphere of social norms that facilitate our joint interactions. Again, these answers must

[3] A similar point is developed by Austin (1962, 83-93) concerning performative terms.

be developed and require further exploration. However, I think these uses are not paradigmatic of the type of phenomenon we capture under the term 'folk psychology'.

4 Evaluation, Action and Justification

In the previous chapter, I have tried to motivate the idea that propositional attitude ascriptions are archetypical evaluative expressions by displaying the weaknesses of folk psychological descriptivism. The aim of this section is to present positive motivations supporting the evaluative view. These motivations go from several everyday observations to empirical research. In particular, I concentrate on three ideas: the connection between action and evaluative expressions, the empirical evidence concerning cross-cultural differences in folk psychological explanation, and some evidences in the field of adults' mindreading. Although I am aware of the limits of these arguments, I believe that, taking them together with the previous arguments against descriptivism makes them compelling enough to embrace the idea that propositional attitude ascriptions are evaluations.

Firstly, our everyday encounters reveal a special connection between propositional attitude ascriptions and behavior. This special connection does not only manifest in the obvious cases of rationalization and explanation of behavior, but also, as I discussed before, in the nature of some peculiar cases of disagreement. Another way to appreciate the connection is by attending to the reactive responses of folk psychologists to the incoherence between ascriptions and actions. McGeer (2015) presents this phenomenon as follows:

> we human beings are deeply invested, both practically and emotionally, in regulating one another's thought and action even outside the moral domain. Consider, for instance, how hot under the collar we can get about people who profess certain beliefs and yet do things that are completely at odds with such beliefs – that is, individuals

> who we disdain as "hypocrites", notwithstanding the fact that we often instantiate that property ourselves. (McGeer, 2015, 171).

The reaction to these kinds of incoherence reveals that we expect people who avow particular mental states to act as they ought to according to the ascription. Contrary to usual descriptivist expressions, one expects the attributees to have the obligation of behaving in particular ways. Nonetheless, these obligations are not at issue when we use descriptive vocabulary. This connection is important when you notice that only evaluative or normative expressions seem to have this practical function. In fact, as I defended previously, the connection is characterized by means of the practicality principle, which Strandberg (2012) introduces in the context of meta-ethics, where the idea that moral discourse is connected to action is widely spread (Hare, 1952, 148-149; 169; Horgan and Timmons, 1992, 164-165; Smith, 1994, 61). In moral discourse, the connection is manifested in the inclinations to act in accordance with your moral expressions. Consider, for instance, how strange would be to say 'torturing animals is wrong' but bid in a dogfight. Other evaluative expressions which possess this special connection with action are deontic modals (Charlow, 2015) and rationality ascriptions (Gibbard, 1990). So, we have reasons to assume that if a particular part of discourse has this special connection with action, then it should be treated as evaluative vocabulary.

The connection between propositional attitude ascriptions and behavior has been specially emphasized in the case of knowledge attribution. For instance, Kappel and Moeller (2013) introduce the point as follows:

> The basic observation is that attributing knowledge to oneself or someone else is systematically connected to being motivated to act in distinct ways. Specifically, when a subject A makes a sincere knowledge attribution of the form "S knows that p", then A will be strongly inclined to stop inquiring whether p and take p for granted in her practical and theoretical deliberations. Just consider the apparent oddity of someone saying sincerely that p is known to be the case, and yet she staunchly continues inquiry whether p, or refuses to

> take p for granted in her practical or theoretical deliberation (Kappel and Moeller, 2013, 1530-1531).

The concept of knowledge has a special connection with behavior manifested in the fact that those who claim to know something regulate and adjust their behavior according to it. The same practical function is exhibited by other propositional attitudes. For instance, consider how strange would be to say 'I am craving for a beer' and then ordering wine when coming into a pub; or imagine the reaction of someone being told 'Sara believes it is freezing outside' but then, seeing Sara walking out in a T-shirt. If the special connection with action could be considered as reliable cue of the evaluative nature of a concept, and propositional attitude ascriptions exhibit this mark of evaluative vocabulary, there is a rationale for considering propositional attitude verbs as evaluative tools for assessment.

Another motivation for the evaluative view has to do with two sources of empirical evidence concerning mindreading. The first one comes from cross-cultural differences concerning folk psychological skills (Lavelle, 2016; Strijbos and de Bruin, 2013b). Part of the rationale behind the received view has to do with the pervasiveness of social coordination in different human populations. Generally, humans are sophisticated social creatures able to tackle social projects which demand social coordination. It is part of the received view that this coordination could not be displayed without the capacity for describing/attributing behavior-causing inner states to the subjects. However, there are several studies that press against this assumption. Firstly, folk psychological concepts, including 'belief' and 'desire' might not be universal across human populations. Some authors (Lebra, 1993; Rosaldo, 1980; Vinden, 1996; Wierzbicka, 2006) have argued that some folk psychological concepts might be culture-dependent. For instance, Wierzbicka (2006) argues that the distinction between 'believe' and 'think', which reflects different degrees of credence, is id-

iosyncratic of Anglo-Saxon cultures. Similarly, Lebra (1993) considers that the Japanese concept 'Kokoro' does not have an equivalent in other cultures[4]. Secondly, some studies detected cross-cultural differences in explanatory strategies. For instance, whereas Chinese, Japanese and Indians tend to explain behavior in terms of the situation or social rules, Americans tend to do it in terms of mental states (Dweck et al., 1995; Miller, 1984; Morris and Peng, 1994; Naito and Koyama, 2006). Morris and Peng (1994) made this point by contrasting the language used to account for two murders in both Chinese and American newspapers. Chinese newspapers accounted for a case of a student who shot his advisor. The chronicle emphasized the situation of the student (he lost an award competition and failed to get a post), the relation with the advisor (he did not get along with his advisor) or social circumstances (Chinese competitive academic environment). On the contrary, the American newspaper accounting for an unemployed worker who shot his supervisor emphasized the psychological states of the subject (very bad temper, unstable)[5]. This example incarnates the well-documented differences between East Asian and American/European ways of thought (Nisbett, 2004) and it reflects the differences at the level of folk psychological explanations.

This evidence raises some interesting points related to the evaluative view. According to this view, propositional attitude ascriptions are instruments of evaluation we display as regulative responses or strategies in particular situ-

[4]The term 'kokoro' is usually translated as mind, will, sentiment or spirit, but none of these terms seems to be an appropriate translation; since, as Lebra (1993) indicates, these terms do not capture the embodied aspect of the concept. Kokoro seems to share features of both emotional and rational aspects of the mind that Westerners usually dissociate.

[5]One may consider that these differences reflect the distinction between situation and trait characters (Jones and Nisbett, 1971). This distinction has been consistently criticized (Malle, 1999; Malle et al., 2007). However, an interesting point regarding these issues is the use of these types of mental states instead of propositional attitude and reasons. I think this is a reflection of a phenomenon I discussed in the previous section, the immorality of the action forces the attributer to avoid reasons that could be understood as a justification. In any case, my point here is more general, emphasizing the cross-cultural differences one may expect from the framework I canvass in this dissertation.

ations. Having said that, one may expect to find cross-cultural variations in both the different strategies particular cultures display and the different explanatory/justificatory instruments they exploit. For instance, one may expect certain cultures to avoid some strategies of explanation and prediction, as some Pacific cultures avoid talking or speculating about what others' persons have in mind (Robbins and Rumsey, 2008; Schieffelin, 2008). Also, one may expect that different cultures exhibit different explanatory strategies. This variability is reflected in the differences between explanations in East Asians and Europeans, but more examples can be found. For instance, as Wierzbicka (2006) has persuasively argued, Anglo-Saxon cultures use substantially more mental terms in comparison with other European languages[6].

There is also a second interesting point concerning such cultural variations. Apart from expecting different types of explanations, one may expect mental vocabulary to vary across cultures and languages. Evaluating someone, I maintained, is placing that person as someone from whom you can expect to behave according to a set of normative structures that prescribe what to do in each situation. You make the person responsible for regulating her behavior according to the normative structures and behavioral patterns that follow from the content. With this insight on board, one can expect different mental vocabularies to reflect cultural differences in the normative structures and patterns, but also, different ways to reflect similar action-guiding information under mental terms. These differences are exemplified by the term 'Kokoro' which captures patterns that Europeans/Americans would associate with different terms (both

[6]Wierzbicka and other linguists provide different numbers that support this idea. For instance, in the London-Lund Corpus of English appears an average of 35 occurrences of I think per 10,000 words (in the Cobuild data for "UK spoken" appears 30 per 10,000 words) Aijmer (1997), whereas in the Swedish database, the average is 2.6 occurrences per 10.000 words of the 'jag tror' (I believe). Similarly, Simon-Vandenbergen (1998, 311) indicates that the Dutch expression 'ik denk' appears 9 times per 10,000 words. Finally, the German Mannheim corpus reflects an appearance of the terms 'ich glaube' ('I believe'), 'ich meine' ('I think/mean'), and 'ich denke' ('I think') which are respectively, 5, 3, and 0.6 per 10,000 words.

emotional and rational). Another way to make the point is to consider similar cultures that otherwise do not have equivalent propositional attitude ascriptions. For instance, Spanish-speaking people have a propositional attitude verb 'esperar' which can be translated into English both as 'hope' or 'expect'. Now, while the evaluative view can explain the differences in terms of different ways to evaluate according to different commitments, the descriptivist must assume the bizarre Whorfian conclusion that Spanish and English speakers do have different mental states we describe when ascribing propositional attitudes in social situation; i.e., when a English native speaker says 'I hope that P' and a Spanish native speaker says 'yo espero que P' they are describing completely different mental states. Be that as it may, cross cultural variation in folk psychological strategies and concepts seems to reinforce some of the virtues of the evaluative view.

Finally, another important source of evidence supporting an evaluative approach concerns mindreading capacities in adults (Apperly, 2011, 2013). Apperly has convincingly argued for the importance of systematically studying adults' mindreading skills. Part of the motivation lies in the worry that empirical studies mostly concentrate on infants, children and non-human animals because authors take for granted that adults are experts in mindreading. However, as Apperly himself and others have demonstrated, this assumption is ill-motivated. Firstly, belief and desire ascriptions in adults are not automatic and effortless (Apperly et al., 2006, 2008, 2011; Converse et al., 2008; German and Hehman, 2006). For instance, Apperly et al. (2006) presented adult subjects with a task where they had to track the location of an object in a false belief task, then subjects were unexpectedly asked for another person's belief about the object's location. These subjects responded more slowly to unexpected questions. The control conditions showed that in similar tasks, where the subjects were

warned that they should keep track of the belief, they did not answer slower, demonstrating that belief ascriptions are not automatic. Furthermore, different experiments requiring the subjects to hold a belief attribution, and then including it in a piece of reasoning for predicting the action of a character have demonstrated that adults find this task significantly effortful (Apperly et al., 2011; German and Hehman, 2006).

These features of adult reasoning about mental states along with other empirical evidence showing that both children and adults perform better in implicit tasks than explicit tasks (Keysar et al., 2003; Surtees et al., 2012) have been taken as a strong empirical ground favoring the idea that there are two independent mindreading systems (Apperly, 2011; Bohl and van den Bos, 2012; Rakoczy, 2015). However, there are other possible interpretations for these findings; namely, explicit reasoning including propositional attitude ascriptions has a different purpose from these skills exercised in predicting others. The point is that explicit reasoning is not only effortful but that it is significantly problematic in certain cases. For instance, Low et al. (2014) found that only 56% of adults answer correctly to questions about prediction in ambiguous perspective-taking contexts. The reason why those tasks are difficult for adults may lie in the fact that the tasks are designed to test anticipatory capacities. However, if the evaluative view is right and propositional attitude ascriptions are paradigmatically evaluations of persons, one may expect adults to be better in attributing mental states in contexts of justification, explanation, condemnation and so on, rather than in anticipatory contexts. This may explain the results of the studies about adult's mindreading. Belief ascription is effortful because the studies are designed to force the subject to reason about belief with anticipatory purposes while, if I am right, belief reasoning is mostly present in justificatory cases.

This interpretation is compatible with the phenomenon, but it is also coher-

ent with other findings concerning the role of explanation of intentional actions and reasoning in general. Firstly, Bertram Malle's already mentioned work (Korman and Malle, 2016; Malle, 2004, 2011) has proven that people offer more reasons in terms of mental states when they face puzzling actions in contrast to ordinary actions. But also, there is a tendency to provide more reason explanations (in contrast to appeal to causal factors or dispositions) when there is a justificatory or evaluative objective. Furthermore, some authors have argued that the primary function of reasoning is to convince others (Mercier and Sperber, 2011; Norman, 2016). According to this view, reasoning has a social function: to provide arguments for social cover and regulating others' behavior to align it with ours. To argue for this hypothesis, Mercier and Sperber discuss a vast amount of empirical literature in reasoning that seems to support it. For instance, the empirical evidence concerning reasoning failures, like the confirmation bias (Nickerson, 1998), or the Wason selection task (Wason, 1966), seems to confirm that reasoning is not an accurate medium for gaining reliable knowledge. Phenomena such as biased evaluation (Lord et al., 1979) or belief perseverance Ross et al. (1975) support the claim that human reasoning has a social function, that is, to 'devise and evaluate arguments intended to persuade" (Mercier and Sperber, 2011, 57). For example, our tendency to reason in such a way that reinforces our prior beliefs, even when these beliefs are false(Nickerson, 1998), supports the idea that reasoning capacities have a social function rather than an epistemic one. Of course, this does not count as direct evidence in favor of the evaluative view. However, it helps to place the view in a more general framework, where reasoning is a tool for regulating behavior rather than gaining knowledge. According to the evaluative view, our mental concepts, including our epistemic concepts, are tools to evaluate persons with a justificatory/explanatory purpose. In a social environment where people are

able to reason to convince, persuade and justify their actions, it makes sense to assume that some conceptual tools that facilitate fine-grained evaluations and specify degrees of support have emerged.

As a result, we have a solid background of motivational forces to embrace the idea that propositional attitude ascriptions are evaluative in nature, that is, they serve to provide significance, credibility, responsibilities, merit, or demerit to a person. They are tools for situating people in a net of duties and rights that helps us to approve or disapprove of them, to regulate their behavior, or to carry out other social objectives. This evaluative view is not only motivated by the problems of the descriptivist contender discussed in chapter 4, but also, by the positive motivations discussed in this chapter.

5 Conclusions

This chapter aimed to provide and motivate a minimal characterization of what is involved in regarding propositional attitude ascriptions as evaluative. According to this characterization, propositional attitude ascriptions are not descriptions, they are evaluations of a person as having different degrees of responsibility, credibility, merit, or demerit towards a particular content. In contexts of folk psychological explanation, this is translated into situating a person as someone you must expect to behave and speak in ways compatible with the normalized patterns that follow from the content. Of course, much more empirical and conceptual research is necessary to elucidate the evaluative character of ascriptions. However, I think that the minimal characterization is substantive enough to motivate an alternative to the theories concerning the relation between language and social cognition presented in chapter 2. In particular, this evaluative alternative explains several aspects of propositional attitude ascriptions that present a challenge for FP-descriptivism. Now, the purpose of

the following chapter is to draw the consequences that this alternative understanding of folk psychological ascriptions have for the debate concerning the role of language in mentalizing. In particular, I will argue that the evaluative view motivates a particular empirical hypothesis concerning the developmental connection between language and folk psychological acquisition, namely, a conversational hypothesis.

Chapter 6

Talking the Way to Other Minds

1 Introduction

The two previous chapters were devoted to defending and characterizing a particular manner of understanding propositional attitudes. According to it, propositional attitude ascriptions are tools for assessing other people in social situations. Paradigmatically, ascription burdens attributees with different degrees of merits or demerits, responsibilities, credit, or significance. Shifting the function of mentalizing from the descriptive focus to the evaluative one must have, one may expect, several conceptual and empirical consequences. This section aims to discuss the main consequence of the evaluative view for the debate concerning the relation between language and folk psychology: the conversational hypothesis. According to this hypothesis, children need to be exposed to particular conversational situations in order to gain the ability to depict other agents in terms of beliefs, desires and other propositional attitudes.

In the following section, I articulate the connection between considering ascriptions as tools for assessment to the idea that conversationally mediated joint activities are the type of circumstances that provides a social scaffolding for folk psychological capacities. After that, I characterize my version of the hypothesis by contrasting it with other similar positions in the literature. Then, I discuss how the hypothesis accounts for some of the evidence presented in chapter 2 and other types of evidence in developmental psychology. Finally, I consider a general objection against my proposal, namely, the possibility that non-human animals possess mentalizing abilities, which could be interpreted as a counter-example to the conversational hypothesis.

2 Evaluation and Conversation

In order to make my point, consider several of the social circumstances where we put mental concepts to work. In the previous chapters I pointed out to several of these circumstances. Firstly, those of justifying counter-normative behaviors. When the normative structures governing our social games make behavioral deviations salient, the situation can be handled by ascribing mental states and reasons that normalize the action. Secondly, in contexts where social agents avoid public sanctions or potential objections, one may avow a mental state to indicate uncertainty or modulate her speech acts for exculpating her action. Thirdly, those of ascribing responsibility to condemn a behavior. Fourthly, contexts where we present our and other's mental states in order to regulate or adjust others' behavior, for instance, when we ascribe mental states to a third person to exercise her authority or when we avow ours in order to convince someone. Finally, ascribing propositional attitudes appears in possible conflicting interpretations regarding actions we need to evaluate. From a developmental perspective, one may notice that paradigmatic examples of these

situations can be found in contexts of reciprocal play, caregivers-children interactions, cooperative games, tasks resolutions, and so on. Of course, this list is far from being exhaustive, but these examples should call our attention upon certain interesting features of the relevant circumstances for our investigation.

The first relevant feature is that these are contexts of joint activity, that is, contexts of moment-by-moment coordination of actions. A basic aspect of these activities is that the actions of the agents involved are relevant to each other, which is a necessary component of acquiring the capacity to evaluate other agents. When others' behavior becomes relevant to the achievement of our objectives and cooperative projects, then to burden others with responsibilities, merits and grades of commitments is a useful tool. The second relevant aspect is that these environments are linguistically infected, carrying on these joint activities depends on dialogues that coordinate how the activity unfolds; but also, using proper evaluative tools that help us to influence others' behavior in order to specify the pertinent courses of actions to display the activity. In other words, the relevant situations are those where there is a coordinate action conversationally mediated where participants have to monitor and adjust each other's participation.

From these conclusions we move naturally into the idea that humans need to be trained in conversational contexts with these features in order to acquire the capacity to attribute propositional attitudes. Acquiring mental concepts involve gaining the ability to evaluate others on the basis of a background of normative structures and making explicit how different situations are linked to certain perceptions, behaviors and speech acts. The paradigmatic environments to earn these skills are those where others' actions become salient, where our actions are exposed to possible sanctions and where our mutual objectives could raise possible conflictive situations that invite us to recognize others' obligations and duties

given the social interaction. From a developmental point of view, those contexts include playing with siblings, friends and caregivers, solving problems with others, pretending cooperative games, cooperative tasks resolution and so on. For instance, consider children tackling cooperative pretend play. In those contexts, children need to discuss a pretend plan or game and to define and negotiate the roles through narratives. This implies situating the different roles, making explicit the goals and characteristics of the roles. During the game, the children's conversations require solving disputes, performing the goals of the game, following pretended suggestions and other actions which require exercising the paradigmatic folk psychological tools for justifying, solving disagreements, regulating each other minds, particularizing goals or exculpating and condemning actions (for empirical support see sect. 4 below). Joint activities involving these types of situations provide the appropriate developmental scaffolding to earn folk psychological skills. In other words, children need to participate and being exposed to conversations of a particular type. The characterization of this empirical hypothesis is the aim of the following section.

3 The Conversational Hypothesis

Now, the question is how to give shape to the idea that conversational environments of those types are the developmental key to acquire folk psychological capacities. Certainly, I am not the first to propose a conversational approach to the development of folk psychology. Other psychologists and philosophers have emphasized the role of conversation in the acquisition and improvement of socio-cognitive capacities (Carpendale and Lewis, 2004; Harris, 2005, 1996; Harris et al., 2005; Cleave and Gauker, 2010) Furthermore, these approaches have been accompanied with a substantive number of studies (e.g. Astington and Jenkins, 1999; Dunn et al., 1991; Hughes and Dunn, 1998; Meins et al.,

2006). In order to introduce my own interpretation of the hypothesis, I would like to contrast it with other two versions developed by Harris (1996, 2005) and Gauker (Gauker, 2003; Cleave and Gauker, 2010) which are conceptually and empirically well supported.

According to Harris (1996, 2005), folk psychological understanding requires the recognition of others' perspectives or points of view. The main feature of mentalizing is to appreciate that others have a different perspective on reality. This feature of mentalizing, Harris argues, is not acquired through syntax or semantics itself but through conversation in general. By being exposed to different exchanges of information, children are reminded that speakers have different beliefs, desires and intentions. For instance, the exposition to conversational exchanges helps children to appreciate the different perspective the mothers intend to introduce the children: "it is the mother's pragmatic intent, notably her efforts to introduce varying points of view into a given conversation, that is the underlying and effective source of variation" (Harris, 2005, 77). Harris demands folk psychologists to acquire the sensitivity to understand others as recipients of information, what he calls epistemic subjects with a particular point of view. What Harris seems to have in mind is the idea that conversations offer to the children the opportunity of engaging in circumstances where they can be forced to engage in simulation to understand others' perspectives or points of view. However, if understanding points of views in simulationist terms requires descriptive-states (section 3) then, Harris' version of the conversational hypothesis is subject to the same type of criticism I have developed in chapter 4.

Closer to my position, we can find the version maintained by Cleave and Gauker (2010). As I discussed in chapter 4, van Cleave and Gauker (see also Gauker, 2003) start with the idea that the primary function of belief and desire ascriptions is to perform vicarious speech acts, i.e., asserting and commanding

on the behalf of other persons. These vicarious speech acts are the basic function of ascriptions that, after that, can be recruited for explaining behavior. In fact, van Cleave and Gauker claim that children can provide partial explanations based on these uses before they engage in complete explanation with reference to mental states. My convergence point with van Cleave and Gauker has to do with the preconditions they impose for belief and desire acquisition. According to their view, the main requirement to acquire the capacity for ascribing beliefs and desires is the recognition of credibility and authority. If the primary function of belief ascriptions is to talk on the other's behalf, then belief attribution requires identifying others as credible spokespersons. Similarly, if the primary function of desire ascriptions is to command on the other's behalf, desire ascriptions require to identify authority. They explain the idea as follows:

> [Belief ascriptions] require an understanding of the practice of contributing information to a cooperative plan. Not only that, but they require an understanding of relations of credibility between one person and another. The attributor needs to understand that, while he or she may not be entirely credible for his or her audience, somebody else may be more credible for the audience, or that while a person is not credible for himself or herself, that person may be credible for someone else. (Cleave and Gauker, 2010, 318).

Recognizing others as being credible or having authority is the key to produce vicarious speech acts, and therefore, to acquire the primary capacities for ascribing propositional attitudes. Although van Cleave and Gauker do not elaborate upon which types of conversational environments are required, they point out to the contexts of dispute as one example of the conversations that help to exercise these primary uses of belief and desire. In fact, they consider the research of Bartsch and Wellman (1995, chapter 6) as evidence in support of their position. In these studies, Bartsch and Wellman discovered that early uses of belief and desire seem to be bound to situations of dissolving disputes, rather than explaining actions. These uses, they argue, can often be viewed as indirect discourse attributions or paraphrasing attributions: "Interlocutors may assert 'You think

that p' and 'I think that q' as a way of defining a conflict in opinions before attempting to resolve it. The same might be said about command- conveying and need-conveying attributions of desire: They may often occur in the context of a conflict of goals as a means of identifying a conflict in goals before attempting to resolve it" (Cleave and Gauker, 2010, 315).

The differences between van Cleave and Gauker's view and my own concerning ascriptions yield differences in our way to approach the conversational hypothesis. Firstly, there is a question of emphasis about what is going on in vicarious speech act cases. They claim that talking on the others' behalf requires the recognition of a certain quasi-legal position, i.e., authority and credibility; while my view maintain that ascribing beliefs and desires amounts to give voice to this position[1] (credibility and authority, but also responsibility, merit, demerit, etc.). In other words, the recognition of authority, credibility or other social status is not a prerequisite to ascribe mental states: instead, it is the primary function of ascribing mental states. When ascribing a mental state, we are giving voice to the credibility, authority or responsibility of this person. Furthermore, vicarious speech acts are not the primary actions that open the developmental path toward descriptive functions for explanation and prediction. On the contrary, vicarious speech acts are only one of the many contexts where we display our evaluations, and this function is shared with paradigmatic explanatory and justificatory contexts. This does not mean that vicarious speech act uses, parenthetical or others, do not appear before in development; my claim is that there is not a difference in function between these uses and paradigmatic folk psychological explanation[2] (i.e., all have an evaluative function).

[1] Furthermore, as I argued in chapter 4, ascribing a belief is evaluating someone as being responsible, which most of the times implies the attributer is not committed to the content expressed by the that-clause. This seems to be quite the opposite to considering someone as credible.

[2] A different matter is whether or not propositional attitude ascriptions can be used with non-evaluative (descriptive) functions (see chapters 5).

Now, contrary to these two versions of the conversational hypothesis, the version I canvass in this dissertation can be introduced as follows:

> **The Evaluative Conversational Hypothesis**: The capacity for evaluating others in terms of propositional attitudes requires the folk psychologist to engage in conversational contexts embedded in cooperative projects and joint activities. In those contexts, the courses of action of the different parts become salient and relevant to each other in such a way that the participants need to evaluate and situate each other in a background of supportive attitudes to norms.

There are two primary differences between the hypothesis I canvass here and the other two versions. Firstly, my version of the conversational hypothesis is practical. Learning to ascribe propositional attitudes is not a question of comprehending others as having a special feature or psychological reality. On the contrary, ascribing propositional attitude is an ability, the ability of evaluating others on the basis of different normative structures, routines and social situations. In this sense, conversational environments are schooling environments that provide a social scaffolding to engage in evaluative practice, rather than special sources of information. Furthermore, as an ability, acquiring folk psychological skills is not only a matter of being exposed to the appropriate contexts, but participating actively in the conversational contexts. Secondly, not all conversational contexts are equally relevant. Giving that acquiring folk psychological skills requires being able to evaluate a person for the connection with action, children need to engage in conversational contexts where their actions are open to possible sanctions; but also, where their mutual objectives could raise situations that invite them to recognize different responses and different obligations and duties given social interactions. In other words, the appropriate conversational situations are those where the children are embedded in cooperative projects and joint activities with others. As I mentioned before, paradigmatic examples of these contexts include: cooperative games, joint pre-

tend play, cooperative tasks resolution and so on. One may be concerned about the fact that once you have embraced an evaluative conception of propositional attitude ascriptions, there are other versions of the views presented in chapter 2 that could be presented under an evaluative clothing. For instance, one may put forward an evaluative version of the semantic view, according to which acquiring the competence to evaluate a person requires being competent in the use of mental concepts and being exposed to them. As I argued in chapter 5, this capacity is framed into a larger conception of folk psychological explanation as an evaluative practice that does not necessarily involve mental state ascriptions (instead it may involve factive reasons, for instance). Thus, I see the conversational hypothesis as including an evaluative version of the semantic approach. Propositional attitude ascriptions require mastering and being acquainted with different uses of mental concepts.

4 Empirical Evidence: First Gear

The rest of this chapter is devoted to discuss different empirical studies in order to motivate the empirical plausibility of the view. In chapter 2, I presented four empirical sources of evidence that helped me to calibrate the empirical hypothesis concerning the developmental relation between language and mentalizing should tackle. I will discuss how my version of the conversational hypothesis –what we can call, generally speaking, an evaluative version of the view– can account for them. Furthermore, they will help me to make clear the particularities of the evaluative version of the conversational hypothesis as opposed to the other two.

4.1 The Appearance of Mindreading

In chapter 2, I discussed three sets of relevant empirical results concerning the appearance of mindreading. These results situate at different ages the time when children could be credited with full-fledged mindreading capacities: the implicit FB-tests (situating mentalizing at 15 months old), the explicit FB-tests (around 4 years old) and the intensional tests (around 6 years old). As a reminder, in the explicit task, a child is presented a scenario where a character, Maxi, puts chocolate into a cupboard x. When Maxi is not present, his mother displaces the chocolate from x into cupboard y. Then, the experimenters ask the children where Maxi will look for the chocolate. The rationale behind the experiment is that only when the child is able to represent Maxi's wrong belief, he is able to point correctly to box x, something that happens around age 4. In the implicit task, the children were exposed to a similar scenario. Then, the experimenters measured the looking time of infants to test their reactions. In this case, the infants look longer when the person, who was not present when the object was relocated, picks the object, something that locates understanding of other minds earlier than it was supposed, around 15 months of age. Finally, in the intensional tests, children were presented with a story or premise where an object has two aspects (A and B). A person does not know that they are the same object and she is looking at it under the aspect A. Then, the children were asked whether or not the person is looking at B (the correct answer was no). Children between 6 and 8 years old find the task difficult to understand. These findings are used to locate the understanding of others' mind later than it was supposed. To add fuel to the fire, an increasing number of cross-cultural experiments demonstrate that the developmental trajectory in folk psychological acquisition may vary dramatically from culture to culture (Liu et al., 2008; Naito and Koyama, 2006; Vinden, 1996, 2002) situating the success in explicit FB-task at different ages

depending on the culture. For instance, Hong Kong children pass the test around 5 years old, while Canadian do it at 3 years old and the Junin Quechua children in Perú do not do it until 8.

Now, the question is how the conversational hypothesis accounts for the different tasks testing socio-cognitive capacities in children. Regarding implicit FB-task, the defenders of the paradigm take these findings as supporting the idea that humans are equipped with an innate mindreading capacity. Before considering this issue, we must notice that these tasks are testing anticipatory capacities. In chapter 4, I have maintained that cases of anticipation based on mental states ascriptions are limited. So, my inclination is to consider implicit FB-tasks as testing anticipatory capacities which do not necessitate mental state ascriptions. On my view, anticipatory capacities are based on normative structures and generalizations that help to generate expectations about what agents ought to do. Given the FB-tasks situations, children may exploit the surrounding circumstances to extract the normative consequence of it. This would make the evaluative conversational views compatible with deflationist interpretations of the implicit FB-tasks (Apperly, 2011; Ruffman and Perner, 2005; Rakoczy, 2015; Zawidzki, 2013) According to which implicit FB-tasks do not test mental state ascriptions but low-level anticipatory capacities relying on situational cues or generalized knowledge acquired through statistical learning capacities. One may interpret the results of the implicit tasks as exhibiting the capacity of children for exploiting generalizations concerning the goals and rational means of the agents in particular circumstances. In chapter 4, I mentioned Zawidzki's view concerning this point, but there are other ways to come out with the same idea. For instance, Ruffman and Perner (2005) contend that purely associative or statistical learning mechanisms can account for the results. Discussing the findings of Onishi and Baillargeon, they say:

> we acknowledge their [Onishi and Baillargeon's] suggestion that infants expect the observed person to act in a particular way. However, we propose that this can be based on behavior rules. Infants may have noticed (or are innately predisposed to assume) that people

> look for an object where they last saw it and not necessarily where the object actually is. Again, such a rule captures something implicit about the mind, because the rule only applies as a result of the mind mediating between seeing and acting. Nonetheless, infants can simply know the rule without any conception that the mind is the mediator (Ruffman and Perner, 2005, 215).

Children's anticipatory looking behavior may rely on rule-based knowledge concerning what agents ought to do depending on their situations and what they have access to. This interpretation can be understood in terms of the normative structures that regulate our social environment[3].

Explicit versions of the FB-task raise some significant methodological issues that concern our investigation. Again, the main problem is that whereas I have defended that propositional attitude ascriptions are mostly justificatory/explanatory tools, these tests are primarily designed to tap anticipatory capacities. There are several things that deserve mention at this point. One may consider that explicit FB-tasks may count as a counter-example against my position because the tests demonstrate that children anticipate behavior by reasoning about others' mental states. For instance, one may be able to say that Maxi will look for the chocolate into the cupboard x, even though his mother moved it to another cupboard y, because Maxi has a false belief concerning the location of the chocolate. However, this is not necessarily the case. As the nativist approaches have convincingly argued (Rubio-Fernández and Geurts, 2013; Westra and Carruthers, 2017; Westra, forthcoming), there are reasons to believe that the questions presented in explicit FB-tasks (where will the protagonist look for an object?) can be misunderstood by the subjects of the experiment. In fact, simplified versions of the task are designed to tap children's understanding before the 4th year of life (Rubio-Fernández and

[3]One could align my view with certain type of constructivism (see for instance Westra and Carruthers, 2017). However, although I consider most of the socio-cognitive capacities to involve exploiting regulative practices that help to learn the permitted behavioral patterns, I do not rule out the possibility that some of the heuristics we exploit to anticipate behavior are innate or rely on important innate mental structures.

Geurts, 2013). Giving that, one may consider that these simplified versions of the tasks are assessing the same anticipatory capacities as the implicit-version of the task. Children's struggles with the task rely on their incapacity to articulate their responses verbally (Rhodes and Brandone, 2014). Then, once they are able to articulate this information, they can verbalize the anticipatory information to solve the task and answer where the subjects will look. This deflationist interpretation coheres with approaches to anticipatory capacities I mentioned in chapter 4, normative anticipation based on rationality constraints or anticipation generated by generalizations about everyday situations. As Andrews (2014)puts it, children could be familiarized with everyday situations, for instance, "Mom looks for her keys where she left them every time she leaves the house, and even if the toddler got her hands on them, Mom still reaches for her keys on the empty hook" Andrews (2014); and then, they can project this knowledge to other situations. This could explain why kids can engage in prediction of these types of situations without putting mental concepts into work. Other types of examples are paradigmatic cases of deception where an agent uses a signal or action to deceive another individual. Again, this behavior can be understood in terms of familiarity with general patterns of behavior in particular situations. So, the agent can misinform others to trigger the particular normalized action that obtains in a situation. The key is that anticipatory skills are more automatic and parsimonious than what folk psychological descriptivism has supposed. Humans and other animals are able to predict others' behavior when attending to the situation the targets are immersed in because they are familiar with the paradigmatic courses of behavior that obtain given the situation.

Now, does that mean we do not engage in forward reasoning involving mental states to bring out predictions? Some authors claim that our anticipatory

capacities rely on reasoning structures such as: 'if A wants p and believes that doing q will bring about p, then ceteris paribus, A will do q' (Borg, 2007, 7). Imagine someone watching Maxi's situations that predict his behavior in such a way: "Maxi believes the chocolate is in the cupboard x, her mother changed the chocolate from x to y. Maxi did not see his mother changing the chocolate in the cupboard x". Now, there is nothing here to make a case against the evaluative view. As I argued in chapter 5, attributing a mental state is to situate a person as someone from whom you expect to regulate her actions according to the social norms that prescribe what to do giving the content of the mental state. This has certain normative imports (obligations and duties), but also, it generates certain expectations about what the person will do[4]. Furthermore, learning to apply these concepts also entails to learn the conceptual connection with other mental concepts. This explains why we can engage in those types of reasoning to make certain expectations. However, this does not involve abandoning the evaluative view.

Other versions of the task explicitly ask for the mental states of the protagonists (what does the protagonist think?). Of course, children need to have an understanding of the concept 'think' to answer this question. However, children do not seem to be asked for an anticipation in this case. On the contrary, they seem to be asked for inferring what the protagonist believes or thinks given the situation. This can be reinterpreted in evaluative terms: which responsibilities we can expect from the protagonist given the situation. Thus, these different questions may test different capacities. Regarding this issue there is some evidence supporting that children are able to pass explanatory versions of the explicit FB-tasks before they can deal with anticipatory versions (Bartsch

[4]Notice again how difficult it would be to make accurate predictions about the behavior of the agent in real life. Maxi could be on diet; Maxi could detest chocolate or be allergic to it; he could forget were the chocolate was or just notice her mother's half-smile when intending to kid him. These conditional situations, captured by the ceteris paribus clause, are what make prediction in terms of mental states intractable (Chapter 4).

and Wellman, 1989; Robinson and Mitchell, 1995). For instance, Bartsch and Wellman presented 3 years-old children, 4 years-old children and adults with different short descriptive theories they had to explain: "Jane is searching for her kitten under the piano. The kitten is hidden under the chair; why do you think she is searching there?" Then, the different answers were analyzed. They found that 3 years old were able to explain anomalous behavior in terms of false belief even when they failed to accomplish predictive versions of the task. These findings do not only support the interpretation that predictive explicit FB-task results are problematic, but it reinforces the idea behind the evaluative view, according to which mental concept acquisition is tied to justificatory/explanatory situations.

Taking these findings together, we can fairly argue that anticipatory and justificatory/explanatory capacities follow different developmental paths. Anticipatory capacities may appear before linguistic abilities develop, then around the third or the fourth year, children start to verbalize this predictive information, but also, they acquire the capacity of providing explanations and justifications in terms of mental states. The fact that these capacities appear in a short period of time, and the methodological decision of taking explanation and prediction as an integrated (see Andrews, 2012, 37-44; for a critic) capacity may have masked the fact that they follow different developmental paths. This is reflected in the explanatory versions of the FB-tasks, but also, in the problems that adults encounter with some explicit anticipatory tasks (see chapter 5). Anticipatory capacities may exploit several oughts concerning the normative import of the situation and social norms. This information is later integrated with mental concepts and reasons structures. That is the reason why adults can use these reasons and ascriptions with anticipatory purposes. Of course, all these conclusions must be taken with caution pending confirmation of more em-

pirical findings. However, I think we have enough arguments to conclude that the evaluative conversational hypothesis can give a plausible account of several results concerning both implicit and explicit version of the FB-tasks.

Finally, the intensional tests may present a challenge to the evaluative conversational view as far as children cannot recognize intensional aspects while they can pass explanatory versions of the FB-tasks. Intensionality is a central feature of propositional attitude ascriptions, so one may expect children who pass the explanatory versions of the FB-tasks to understand others as having different point of views concerning a particular event. The problem with these experiments is that, as Rakoczy (2015) has emphasized, "these types of tasks arguably involve performance factors other than what they aim to measure (understanding aspectuality), linguistic demands in the first type of task, and other cognitive factors in the latter. In particular, these tasks involve ambiguity of referential expressions ("an A") and therefore the necessity for reference resolution (which A?)" (p. 5). In order to make his point, Rakoczy et al. (2015) designed a simpler version of the task which does not involve such ambiguity. In this simplified version, the subjects were presented with an object with two aspects (A and B). There were three steps in the experiment: (1) The object was presented under the aspect A and put it in a box 1 in presence of a figure (Peter). (2) The object was presented under its aspect B and put back into box 1. (3) The object was moved to box 2 in front of Peter. The test question is where Peter will look for the object. Then, the subjects were asked where Peter will look for A (when Peter was not present at the step 2). Rakoczy and colleagues found that once the ambiguity disappears, children understand intensionality around the same age they can pass standard FB-tasks (4 years old). In principle, this is expectable from the conversational hypothesis because we can expect children to be able to pass explanatory versions of the FB-tasks

to understand intensionality. Sadly, the experiments only include anticipatory versions of the false belief task. Thus, we cannot know whether the intensional test is correlated to explanation and prediction.

Summing up, the conversational hypothesis predicts that children acquire mentalizing when they are competent speakers able to engage in conversation exchanges where others' actions become salient. These skills connect folk psychological abilities to language. However, as the cross-cultural experimental results indicate, socio-cognitive skills can appear at different ages depending on the idiosyncratic cultural patterns of regulative responses displayed to solve counter-normative actions. Finally, there is another methodological issue that concerns the evaluative conversational hypothesis at this point. In the previous chapters, I have argued that explanatory responses are paradigmatically triggered in contexts of counter-normative behavior or when the expectations of the attributee are violated. The experimental research of Korman and Malle (2016) points out in that direction, but these contexts are rarely tested in developmental psychology. An exception to this is the research of Legare et al. (2010). In these studies, subjects were presented with events that were both consistent and inconsistent with their previous knowledge. Legare and her colleagues found that inconsistent events systematically prompt more explanation than consistent events. Although more studies in this direction are necessary, these results reinforce the idea that explanatory/justificatory strategies are reactive responses to violations of expectations.

4.2 Vocabulary, Perspective-shifting and Syntax

In Chapter 2, I discussed several studies relating folk psychological skills and complex linguistic abilities. For instance, those connecting social understanding to the exposition to mental vocabulary and to children's linguistic competence

(Astington and Jenkins, 1999; Cutting and Dunn, 1999; Happé, 1995; Jenkins and Astington, 1996; Ruffman et al., 2002; Watson et al., 2001). These experiments strengthen the conversational hypothesis: maintaining a coherent psychological discourse facilitates the acquisition of folk psychological understanding because, through those dialogues, children are familiarized with the evaluative function that helps them to evaluate others' reasons for actions in order to adjust and regulate actions.

The other two sets of experiments I reviewed in chapter 2 were those showing a correlation between some perspective-shifting discourse and syntactical properties, and success in FB-tasks. On the one hand, Lohmann and Tomasello (2003) designed an experiment where subjects were presented with a deceptive object (a pen that looks like a flower) in different ways: using mental and communicative verbs, using linguistic structures with complementation, using linguistic structures without complementation, and using no linguistic expressions. The subjects of the experiments improve significantly in all training instances except in no language training. The interest of these studies lies in the importance of the perspective shifting discourse:

> The difference between the no language training condition and the discourse condition suggests that in our training this was the explicit labeling of the speaker's perspective. In the discourse only condition the two perspectives were encoded in contentful linguistic symbols, such as, "First it is a flower, and now it is a pen." Therefore, the effect of language really had to do with the adult using conventionalized symbols (mainly in the form of common nouns) to highlight the differing perspectives (Lohmann and Tomasello, 2003, 1139).

Training in a task where two labels can refer to the same object facilitates the subjects' recognition of different perspectives toward an event. On the other hand, the experiments concerning syntax demonstrate the connection between complementation acquisition and success in the false-belief task(De Villiers and Pyers, 2002; de Villiers, 2005; De Villiers and De Villiers, 2000; Hale and Tager-Flusberg, 2003). In the experiments, the researcher measures the success of

the subjects in linguistic tasks including sentential complements. Children who answer this type of questions correctly succeed in the false-belief task more often than children who do not. Likewise, Hale and Tager-Flusberg (2003) found that training children in sentential complement exercises improves their scores in false belief tasks.

In order to see the implication of these experiments for the conversational hypothesis, let me unpack what I think these experiments test. What they seem to share is that linguistic capacities allow to represent a particular event from different perspectives. From the evaluative view, this means that linguistic capacities allow us to engage with the normative significance of the events represented by a particular expression. In other words, different events trigger different behavioral responses. Arguably, perspective-shifting discourse training makes children aware of the fact that a particular object can be represented in different ways. This opens the possibility that a particular event could have different normative import depending on the aspect one is highlighting. From the conversational hypothesis, these results reflect a component of the appropriate situations that help children to recognize how the same event have different normative imports depending on different aspects, and therefore, it could play different reasoning roles depending on the significance that the event has for the person we evaluate.

Regarding syntax, one may admit that folk psychological evaluations in terms of mental state have certain idiosyncratic syntactic features. In particular, mental verbs have the capacity of ranging propositions[5]. Of course, acquiring the capacity to use propositional attitude verbs require mastering these syntactical features. However, I have doubt that it is this isolated capacity what the experiments are really testing (see also Gomila, 2012, 87). My doubt comes

[5]From the logic-syntactical point of view, mental state verbs can be regarded as second-order operators. To see how this feature is related to the evaluative power of mental states predicates see Frápolli and Villanueva Fernández (2012).

from the fact that the training tasks involved in the experiments cannot filter a factor that is important from the view I am canvassing here, namely, the justificatory structure of reasons. In the training experiments, the children were drilled in a task where the character did some action toward a Sesame character but he says he did it to another. Then they were asked to report 'What did he say?' (Hale and Tager-Flusberg, 2003). The trials were designed to highlight the counter-normative behavior of the character. The character says the contrary of what he did. These contexts have the basic structure of the situation that would trigger the justificatory responses that the conversational hypothesis would predict. Obviously, it is also doubtful that this component can be isolated in the longitudinal studies.

As a result, both sets of experiments reflect the importance of linguistic training for recognizing the normative import of events for evaluating behavior. According to the conversational hypothesis, linguistic environments school children in the ability to evaluate other persons and their behavior according to the grades of commitments and responsibilities they exhibited toward particular events and values that permit and prohibit certain behavioral patterns. In this sense, the two sets of experimental data reveal two necessary aspects for exercising this capacity. Firstly, recognizing that the normative import of a particular event would depend on different aspects of the event under consideration. Secondly, being exposed to cases of counter-normative behaviors that facilitate the necessity of evaluating the actions and of searching for justifications/explanations that normalize those patterns.

5 Empirical Evidence: At full throttle

In the previous section, I attempted to show how the conversational hypothesis could account for some evidence concerning the impact of language on social

cognition. Given the methodological issues raised above and the fact that some of those experiments focus on aspects that do not necessarily track the relevant features of language that the conversational hypothesis highlights, I must recognize the restricted impact of those data for tipping the scale in favor of the conversational hypothesis. Of course, this is an empirical hypothesis that needs the development and design of experiments that goes beyond the purpose of this dissertation and the competence of the writer. However, I think there are solid empirical results that point in the direction of the general ideas I have defended in this dissertation. This section aims to review some studies that would help to lay an empirical bedrock for the conversational view.

The first empirical findings that provide some grip to the conversational hypothesis have to do with the quantity of appropriate conversational contexts children encounter. For instance, one may expect that children with siblings confront a more significant quantity of appropriate conversational contexts than those who do not have siblings. A series of studies supporting this idea have found more success on false beliefs tasks in children from larger families (Jenkins and Astington, 1996; Lewis et al., 1996; Perner et al., 1994; Ruffman et al., 1998). In particular, Perner et al. (1994) for instance, found that the number of siblings correlates positively with the score of children in false belief tasks. Furthermore, they found that the second born of two scored better than the first born; and the third born of three was better than the second born, which seems to block the interpretation of maturity as the relevant component.

An interesting point concerning my version of the conversational hypothesis has to do with the appropriate type of conversational circumstances that function as a scaffolding for acquiring propositional attitude ascriptions. As I emphasized before, not all conversational situations contribute equally well to the development of folk psychological understanding; joint activities, where

other's actions are especially relevant for our goals and actions, are the type of social environment that the evaluative conversational hypothesis situate on the basis of the development of social skills. One of the relevant activities I mentioned in the previous sections was cooperative pretend play. Cooperative pretend play, I argued, provides paradigmatic examples of conversational situations whose structure is suitable for facilitating propositional attitude ascriptions. In pretend play, children negotiate the role of the characters, their objectives, reasons and goals, which implies making explicit the responsibilities, commitments and obligations of the characters with particular patterns of actions. A possible alternative interpretation is to consider that is pretend play broadly considered, instead of cooperative pretend play, which is relevant for the acquisition of social skills. However, as Harris (2005) has discussed, there are several reasons to assume that it is cooperative pretend play in the connection with action what makes the difference. For instance, solitary pretend play does not seem to correlate with social understanding (Harris, 2005, 78) while role assignments or the frequency of joint proposal do (Astington and Jenkins, 1999; Jenkins and Astington, 1996).

This idea finds support in the studies carried out by Hughes and Dunn (1997)which show that children are more likely to use mental state terms in the context of pretend and reciprocal play than in other contexts (see also Jenkins and Astington, 2000). These findings were replicated by Hughes et al. (2006). During the studies, the experimenters visited 140 families with 2-years old children and their siblings. The experimenters prompted the subjects to play with costumes and toys that elicited them to engage in pretend play. After those visits, the children were tested in different theory of mind and verbal tasks. Among the results, Hughes et al. (2006) emphasized reciprocal and cooperative pretend play as one of the stronger predictor (along age and verbal skills) of the

use of mental state terms.

Given that the appropriate circumstances involve cases where the actions of the agents are relevant to each other, one may expect situations of prohibiting another's action, protest, threat, insult or refusal of a request; that is, contexts of disagreement and dispute. In this sense, the conversational hypothesis may account for the empirical research emphasizing the importance of dispute or discussions. In the previous section, I presented how van Cleave and Gauker considered the research of Bartsch and Wellman (1995, chapter 6) as evidence in favor of the idea that children start to talk about belief and desire in contexts of resolving disputes. A second set of studies pressing in this direction are those involving the correlation between success in false belief task with argumentation and conflicts (Brown et al., 1996; Foote and Holmes-Lonergan, 2003; Slomkowski and Dunn, 1992). For instance, Foote and Holmes-Lonergan (2003) confirmed two hypotheses concerning this point. Firstly, children's use of mental state terms that occurred during conflicts with siblings was a solid predictor of performance on false-belief tasks. Secondly, children who use other-oriented arguments, which involves negotiation or reasoning that incorporates either the partner's interests or the interests of both children, during these conflicts scored higher in false belief tasks. These findings are important for the conversational hypothesis for two reasons. Firstly, they show the importance of conflict in the development of the understanding of others. Secondly, the procedure to test the folk psychological capacities does not only include anticipatory tasks but also explanatory tasks which include providing reasons about figures' false beliefs. These findings are coherent with other experimental findings set by Dunn and colleagues, who found that children who engage in explanatory discourse in the context of a dispute with their mother and siblings at 36 months were more successful in mental state understanding at 40 months (Dunn and Brown 1993;

Tesla and Dunn 1992; see Dunn, 1994, for a review).

Dispute and cooperative pretend play are frequent between siblings but they are not necessarily the only conversational environments that the conversational hypothesis predicts to correlate with socio-cognitive capacities. Cooperative activities are also important in this respect. In this line, Dunn et al. (1991) found that the degree of children's cooperative interaction with their siblings was related to their performance on various socio-cognitive tasks 7 months later. What all these environments share is that they trigger different regulative responses that facilitate the appearance of explanation and justification of particular behaviors. This would serve as a support for the conversational hypothesis, but there are other studies pointing in the same direction. For instance, Dunn and Munn (1987) studied the appearance of justification in the contexts of discussions with mothers and siblings through a longitudinal study with 43 children. They found that while children around 18 months old tackle dispute mostly with expressions of feelings, they start to provide more justifications based on conventions and the consequences of behavior to deal with dispute around 36 months old. Furthermore, siblings and mothers also increase their justification when solving these conflicts in correlation with the age of the subjects, demonstrating that the expectations of what the subjects understood vary with age. Another remarkable point of the study was that children's justificatory tendencies were more salient in contexts where the subjects' behaviors were prohibited or in cases where the subjects were responding to mothers' justifications. Albeit this evidence is far from conclusive, there is another important consequence supporting the conversational hypothesis. Justificatory skills appear to be reactions to prohibition and mothers' reasons, that is, justifications are bind to the type of situations one may expect when advocating for the evaluative frame.

Looking beyond the propositional attitude ascriptions, a similar connection

between explanations and emotional understanding has been shown by different studies (Denham et al., 1994; Garner et al., 1997). After distinguishing between total reference to emotion (e.g. uses of terms such as 'sad', 'happy', 'anger') and explanations involving emotional terms (e.g. 'I'm angry because you are pulling my shirt') in mother's talk, Denham et al. (1994). found that only the latter affect children's emotional understanding. Garner and colleague's studies deliver similar results. Although emotions are not the same type of mental state as propositional attitudes, there is no reason in principle to suppose that propositional attitude ascriptions may not follow the same developmental pattern.

Finally, one may expect children whose access to these types of conversations has been impeded to have a certain delay in comparison with normal children. In the classical experiments of Peterson and Siegal (2000) with deaf children, the authors found several differences in socio-cognitive capacities between deaf-children who do not have access to conversational interactions during development because of their later signing language acquisition, and deaf native-signing children. These experiments do not necessarily support the conversational hypothesis in contrast to other cognitive views presented in chapter 2; nevertheless, Peterson and Slaughter (2006) designed a set experiments that shows the connection between socio-cognitive comprehension and the skills of children to use mental state expressions in explanation. In these experiments, late-signing deaf children, native-signing deaf children and typical-hearing children were presented with a picture book and purpose-built picture and they had to create a narrative for explaining the pictures. The results, including regression analysis, demonstrate a strong correlation between frequent narrative talk about mental states and score in false belief tasks (including explanation tasks). Another interesting finding of this study is that children's complex narrative

elaboration on mental state terms for explaining behaviors of the characters, or spelling details of their thoughts, were better predictors of socio-cognitive scores than other measured features such as length of narrative, hearing status or general linguistic competences. Besides replicating the studies of Peterson and Siegal, the experiments demonstrate a general delay in socio-cognitive capacities in non-native signing deaf children; but also, that the correlation between narrative explanatory capacities in mental terms is connected to general socio-cognitive capacities independently of the impediment.

Summing up, there is a solid background of experimental studies fostering several aspects of the conversational hypothesis. Generally, there are good indicators of the connection between conversations, linguistic abilities and exposition to general socio-cognitive capacities. Furthermore, there are several studies showing that children's schooling in propositional attitude ascriptions is related to the conversational environments predicted by the conversational hypothesis; namely, joint activities including cooperative pretend play, cooperative actions, disputes and disagreements, and so on.

6 Animal Minds, Language and Folk Psychology

The aim of this section is to encounter an general objection agains the evaluative conversational hypothesis.. This objection takes as a starting point the purported capacity for mindreading in non-human animals (animals from now on). An extended research in animal cognition has tried to elucidate whether or not animals (mostly great apes, dolphins and other cetaceans) can attribute knowledge, perception, beliefs or other mental states to other agents (O'Connell and Dunbar, 2003; Krachun et al., 2009; Tschudin, 2006). Although the possibility of animal mindreading is highly controversial (Heyes, 2014; Povinelli and Vonk, 2003, 2004), one may attempt to exploit the evidence favoring folk

psychological skills in animals against the conversational view. Discussing insightfully the variety of evidence concerning animal social cognition and the different arguments to support different interpretations of these findings is beyond the scope of this section. However, I would like to provide a possible line of response the conversational view could exploit to reply to this argument. In order to face the challenge, I will consider the recent results produced by the studies of Krupenye et al. (2016).

Before targeting the argument, I would like to say a word about the room for animal folk psychology allowed by the view I canvassed in this dissertation. In chapters 4-5, I restricted the use propositional attitude ascriptions to cases where social creatures display reactive responses to counter-normative behavior or other types of circumstances where evaluating a particular agent is required. Furthermore, I argued that propositional attitudes are only a part of the possible reactive responses that social creatures usually deploy. Accepting these two ideas implies that propositional attitude ascriptions and factive reasons are a small part of our socio-cognitive skills. Therefore, the truth of the conversational hypothesis does not erase the possibility that animals exhibit a rich endowment of social capacities. In fact, some of the authors I have aligned myself with have made a significant effort to incorporate animal socio-cognitive capacities into the framework. Following this line, Kristin Andrews has argued that non-linguistic social creatures navigate social interactions by attending to a background of implicit norms governing these interactions. Andrews (2009) starts from the idea that "at least an implicit understanding of normative rules is prior to a theory of mind" (p. 439; see also Chapter 4). After it, she reviews a large set of empirical data in primate cognition demonstrating the understanding of norms in different primate groups. There are three examples of normative sensitivity in primates that are important for Andrews's point. Firstly, relevant studies including chimpanzees, orangutans, and capuchin monkeys demonstrate a substantial variation in culture (Perry et al., 2003; Van Schaik et al., 2003; Whiten

et al., 1999) Among the practices revealing cultural variations, we can find different criteria for sharing food, use of tools and social behavior. For instance, capuchin monkeys exhibit pro-social behavior consisting in various rituals of introduction that enhance trust and bond among members of a population Perry et al. (2003). For example, moving slowly, having one another's fingernails in their nostrils, tail sucking, or hair-passing games. These behaviors, primatologists suggest, are signals of commitments to social relationships. Secondly, in the context of the group, chimpanzees exhibit significant variations in behavior depending on different variables including the number of individuals, whether or not they are hunting in enemy territory, or if they are patrolling their own territory (Mitani et al., 2002). Finally, Boesch (2002) has reported a high variety of hunting strategies and sharing meal rituals in chimpanzees. When hunting, these chimpanzees usually take four different roles: driver, chaser, ambusher and captor. When the prey is selected, each member takes a role depending on the location, the relation to the prey and other members' behavior. Each role has different normalized patterns, which are highly flexible depending on the context. These three examples, Andrews concludes, demonstrate a sophisticated understanding of societal norms and cultural expectation of rule-following:

> While there isn't evidence that nonhuman primates have a theory of mind, they do have the ability to develop variations in their behavioral repertoire that involve creating, following, and violating social norms having to do with trust, harm, and cooperation. These primates appear to have societal norms, and individuals appear to have at least an implicit understanding of the relevant normative rules (Andrews, 2009, 444).

This normative understanding and the capacity to anticipate others' behavior in the light of such norms demonstrates that mentalizing is not required to anticipate others' behavior concerning these norms. In other words, animals can exploit some type of normative strategies to anticipate behavior.

Taking this inside on board, it is possible to encounter the first challenge regarding empirical evidence concerning animal mindreading. Recently, Kru-

penye et al. (2016) have discovered that orangutans, bonobos and chimpanzees can anticipate others' actions in paradigmatic situations of false belief prompted behavior. Using the technology of eye-tracking, they discovered that apes look to the place where an agent must search for a brick that was relocated when the agent was absent. They take these results as proving that apes have the capacity of anticipating others' behavior on the basis of false belief attribution. Now, if these results are compelling, we have reasons to doubt about the connection between linguistic skills and folk psychological understanding that the conversational hypothesis maintains. In order to face the argument, consider the similarity between these results and the findings of implicit false-belief tasks in children. As Heyes (2017) has noticed, there is no way to figure out if we are facing a case of belief ascription or a case of 'submentalizing' capacity that exploits contextual cues in order to interpret visual information. Heyes reinforces this interpretation by appealing to a series of experiments designed by Chun and Jiang (1998) that show how human adults can learn the context of a visual target during visual search in a way such that they can exploit the contextual cues in order to guide subsequent encounters. Assuming similar abilities in apes, one may reinterpret Krupenye et al.'s results as a case of implicit association of the visual context during the familiarization with the task. From this interpretation, apes may have learned to associate the location with the behavior of the target during the familiarization period, so the eye-tracking response is a mere reflect of the expectations generated by the association.

I believe that Heyes' interpretation is too deflationist and fail to consider several important social skills that can be involved in these type of situations. For instance, primates exhibit a certain sensitivity to intentional actions, attention and following gaze (see Gomez 2004, chapter 8; Call and Tomasello 2008, for a review). However, I agree with Heyes in the idea that these experiments do not

require to engage in mental states ascriptions. As I said, following Andrews, apes, as children and adults in implicit versions of the FB-task, may exploit some type of statistical generalizations or normative structures to anticipate behavioral responses (Zawidzki 2013, 15; Ruffman and Perner 2005, 215).

A different way to attack the conversational hypothesis relies on the idea that we can find evaluative abilities of the kind the hypothesis demands in non-linguistic creatures. This argument has been recently put forward by Andrews (2015), who claims that great apes exhibit expressive behavior that can function as tools for justification and exculpation. Andrews' argument is designed to undermine the position of Zawidzki (2013), so let me introduce his view first. Both Zawidzki and Andrews agree about the claim that explanations in social situations are reactive responses to counter-normative behaviors. Ascribing propositional attitudes normalize patterns of action that are perceived as anomalous. However, contrary to Andrews, Zawidzki (2013) tied the appearance of ascriptions to linguistic contexts where two interpretations of a pattern of action conflict. In a population where individuals have both the capacity to express different assertions expressing commitment to normalize their behavior (first-person interpretation) and the capacity to anticipate others' behavior through social norms (third-person interpretation), both interpretations may conflict in some point. Then, propositional attitudes function to rehabilitate the social statuses for avoiding public sanction. Once first-person interpretation may conflict with third-person interpretation of behavior, a distinction between behavioral appearance and mental reality appears on the scene:

> When the interpreters are surrounded by interpretive targets that are constantly making discursive commitments of various kinds and, at the same time, engaging in behavior that may or may not be rationalizable in terms of those commitments, interpreters must inevitably grapple the question: what do they really think? (Zawidzki, 2013, 218).

When the third-person interpretation of a behavior conflicts with the behavioral commitments expressed, we need to appeal to non-obvious mental states in order

to justify o rationalize the behavior, rehabilitating the status of the target after the anomaly. In this sense, propositional attitude ascriptions do not merely express behavioral commitments, but they have an exculpatory function.

Now, Andrews proposes that exculpation can be exercised by other kinds of non-linguistic acts of communication, for instance, pointing, gesticulating or pantomiming. In order to make her case, she appeals to the case of Panbanesha, an enculturated bonobo that responds to a why-question in an exchange with her caregivers. Here is how Andrews explains the exchange:

> In the video Panbanesha and Sue Savage-Rumbaugh are in an informal food context, and Panbanesha is biting on a bowl of food. Savage-Rumbaugh asks "Why are you biting your bowl?" and Panbanesha responds by opening her mouth wide and touching one of her teeth in the back of her mouth. Savage-Rumbaugh, unruffled, says, "I see, your tooth is the problem. Can I look at your tooth?" and Panbanesha allows Savage-Rumbaugh to examine her mouth (Andrews, 2015, 54).

For the purpose of our discussion, the example of Andrews seems to demonstrate that reason explanation does not require linguistic communication. Primate non-linguistic communication, including pantomiming or gesturing, is enough to highlight events or values that normalize behavior. In this sense, Andrews seems to defend that non-verbal communication can work as normalizer of misconducts in the way factive reasons can do it.

There is an aspect of the counterexample that seems quite problematic. It is hard to admit that the case of Panbanesha is a representative case. This bonobo has been exposed to a strong process of enculturation, symbolic language schooling and active teaching training. Thus, she is not a representative sample of non-human cognition. However, my basic concern with Andrew's argument is conceptual rather than methodological. The main point of Zawidzki's view, but also of the conversational hypothesis, is that the concepts of justification or condemn are related to the concept of public commitment. Condemning or justifying someone's behavior requires the attributer to be able to evaluate

the attributee as someone who can be burdened with responsibility and commitments, that is, someone who is in charge of the action. In order to do that, we do not only need to consider a person as someone who follows social norms and state a fact that is normatively connected to the behavior, but to recognize the responsibility or commitments that this person has with regulating his behavior according to these norms. This ability is a side effect of intentional acts of communication which carry, contrary to mere behavior, a public commitment to what follows from the content of the act[6], that is, explicit taking of responsibilities. Now, the question is whether or not these expressions of responsibility are found in animal communication. Paradigmatic cases of non-human communication usually are used to trigger specific responses from the recipient, for instance, warning co-specifics, calling for possible mates, or intimidating competitors. Rather than public acts of commitment, non-human communicative signals trigger specific and inflexible patterns of behaviors. One can see the difference when false signals take place. When an act of communication is false, humans respond by sanctioning behavior, namely, asking for reasons, indicating the falsehood, and so on and so forth. However, for instance, when the capuchin monkeys are exposed to continuous false alarm signals, they just ignore the alarm (Wheeler, 2010). They do not exhibit frustration, punishing actions, or other kinds of responses associated with understanding the communication in terms of public commitments. This points out that there is no understanding of the communicative acts as public signals of behavioral commitments.

Of course, this argument is open to empirical counterexamples. I am aware

[6]One could make this point stronger by arguing that evaluating a person in the way required for exculpating and condemning a behavior does not only necessitate to perform and understand acts of communication that carry those commitments; but also, expressive or linguistic tools that allow the attributer to assign and specify how those commitments must be understood. In other words, from the ontogenetic/phylogenetic point of view, evaluations would demand linguistic devices as the propositional attitude verbs. This strategy would imply to defend that certain evaluations that do not contain evaluative expressions (as factive reasons; see chapter 4 and 5) are subsidiaries, in a sense, of those evaluations which contain them. I will not exploit this possibility here.

of the possibility that future investigations in non-human communication could demonstrate non-human animals can communicate in the way my argument demands. However, these findings, I believe, rather than undermine the argument that evaluations require language, would provide evidence in support of the proto-linguistic non-human communication. Therefore, I would be inclined to admit that animals have the faculty of engaging in folk psychological explanation. This would not undermine the connection between language and folk psychological understanding. On the contrary, I would admit, those non-human animals would enjoy genuine proto-language.

7 Conclusions

The aim of this chapter was to anticipate a developmental hypothesis of the relation between language and mentalizing on the light of the evaluative view. According to this hypothesis, the capacity for mentalizing others –the mastery of attributing propositional attitudes– develops in the context of joint activities mediated by conversations. Notably, these contexts where others' actions became salient, where our actions are exposed to possible sanctions and where our mutual objectives could produce possible conflictive situations that invite us to recognize others' obligations and duties given the social interaction. These contexts include playing with siblings, friends and caregivers, solving problems with others, pretending games, cooperative tasks resolution and so on.

This conversational hypothesis, I attempt to establish, has enough empirical grounding to be considered a plausible developmental hypothesis concerning the relationship between language and folk psychology. This developmental hypothesis was based on a philosophical approach to the main function of folk psychological explanation in general and propositional attitude ascription in particular. Namely, the idea that reason explanation and propositional attitude

ascriptions are not descriptions of psychological states, but evaluations of attributees concerning their responsibilities, merits and significance in connection with their action.

Chapter 7

Conclusions

The primary aim of this dissertation was to offer a hypothesis of the relation between language and folk psychology. The main idea was to defend a type of cognitive view according to which complex linguistic communication is a requirement for acquiring the conceptual capacity of ascribing propositional attitudes. The Evaluative Conversational Hypothesis I canvassed in this dissertation proposes a developmental connection between children's conversational capacities to engage in cooperative actions and their capacity to attribute propositional attitudes. Understanding others as minded comprises evaluating other agents as being responsible for certain contents which are connected to particular patterns of behavior.

There are several developmental findings that endow the hypothesis with a solid background of empirical evidence. In general, there is an important numbers of studies suggesting a strong connection between, on the one side, certain linguistic abilities and exposition to conversations and, on the other, the capacity to ascribe mental states. Furthermore, I have reviewed different studies that make the evaluative conversational hypothesis preferable to other versions

of the conversational view (chapter 6: section 5). In particular, children seem to make use of ascriptions in contexts of dispute and conflicts more often than in other types of conversational contexts. Also, cooperative environments such as pretend play or cooperative games are connected to explanatory capacities and mental state ascriptions. What all these environments share is that they trigger different regulative responses that facilitate the appearance of explanation and justification of particular behaviors.

Beside these developmental finding, the primary rationale behind the evaluative conversational hypothesis was a particular philosophical approach to the nature of propositional attitude ascriptions. According to this approach, propositional attitude ascriptions are evaluations. Propositional attitude ascriptions assess agents as having different levels of responsibility, merit, demerit or significance towards a particular content. Considering propositional attitude as evaluations implies to understand ascriptions as tools for justification, condemnation and rationalization in situations where our actions are permitted or prohibited by a set of norms that govern our social interactions. Having this philosophical analysis in mind would explain why engaging in the conversational contexts explained above connects our mentalizing capacities to complex linguistic communication.

The evaluative view was presented as an alternative to the received view in social cognition according to which propositional attitude ascriptions are descriptions, i.e., representations of internal entities that mirror our internal psychological machinery for the sake of prediction and explanation. The reason to promote the evaluative view over folk psychological descriptivism was due to two different sources of argumentative endowments. On the one hand, folk psychological descriptivism is challenged by a number of problems (chapter 4). Firstly, some uses of propositional attitude ascriptions (parenthetical and communica-

tive) do not have an obvious interpretation from descriptivism and their use reveal an evaluative function. Secondly, folk psychological descriptivism cannot account for a particular phenomenon: the persistence of some types of disagreements. Finally, certain types of explanatory uses of mental state ascriptions are bound to situations where the attributer is presented with counter-normative behaviors. These explanations only makes sense when considering ascriptions as evaluations of the subjects as someone with certain merits or responsibilities that rationalize the counter-normative conduct.

On the other hand, there are several features of propositional attitude ascriptions that allow us to incline the balance to the evaluative side (chapter 5). Firstly, propositional attitude verbs share certain features with other evaluative expressions such as moral terms, for instance, their special connection with action. Secondly, there are certain phenomena one may expect by considering propositional attitudes as evaluations, in contraposition to descriptions, which seem to be shown by some empirical research. For instance, one may expect people to be more skillful using propositional attitude ascriptions in explanatory situations than in predictive situations.

In conclusion, if this research has been persuasive, there are both strong philosophical reasons and empirical evidence to consider the evaluative conversational hypothesis as a serious contender about the relation between language and folk psychology. Furthermore, even if one is not persuaded about the conversational position, I hope to have made a strong case for the necessity of bringing into philosophical focus the assumption behind the empirical debate concerning the role of language in folk psychology. A large number of developmental psychologists, cognitive scientists and philosophers of mind have uncritically maintained a particular approach to the nature of our mental state ascriptions. Thus, even if resisting the evaluative approach, one must face the arguments I

have put forward in this dissertation, or at least, being aware of the necessity of assessing critically the philosophical assumptions behind the different empirical views are based on. A similar conclusion can be drawn for other shared ideas in the debate. For instance, the claim that our social capacities are mostly supported by mechanisms of explanation and prediction.

From these conclusions we can move naturally into how several of the ideas articulated in this dissertation could open different avenues for future researches. From the empirical point of view, the methodological concerns about the different FB-tasks raised in chapter 6 demand the necessity of designing empirical tests able of distinguish anticipatory situations from these of explanation and justification. Further research is required to corroborate or invalidate the evaluative conversational hypothesis. In this sense, it would be necessary to design different experimental setups that help us to decide between the conversational hypothesis and their competitors, but also, between the evaluative version and the others.

From the conceptual point of view, there are different lines of investigation which could start from the ideas presented in this dissertation. Firstly, one may take some of the critical assessments of folk psychological descriptivism as a starting point to reinforce theoretically some of the arguments and hypotheses I have attacked. Secondly, if I am right, and folk psychological descriptivism is a pervasive idea behind the debates in folk psychology, one may exploit the theoretical consequences of considering anti-descriptivism in general, or the evaluative view in particular, in other areas of social cognition research. For instance, the debate about the nature of the mechanisms underlying mental ascriptions. Finally, one may consider to face the two main theses of my dissertation and their connection. Whereas one may reject descriptivism, and embrace the evaluative view, one may doubt that this philosophical approach

necessarily motivates or connects to a conversational hypothesis of the relation between language and folk psychology. In fact, Andrew's arguments analyzed in chapter 6 (section 6) could be interpreted in that way. Admittedly, although I have attempted to give an answer to these arguments, there are several possible replies and findings that would demand further theoretical development of the connection between the evaluative view and the conversational hypothesis. In light of this, I believe it is fair to assert that the philosophical investigation presented in this dissertation would justify several future lines of research, some of which I would like to explore in the following years.

Conclusiones

El objetivo principal de esta tesis era articular una visión cognitiva sobre la relación entre el lenguaje y la psicología popular. Esta visión, que he denominado Hipótesis Conversacional Evaluativa, propone que los agentes sociales necesitan desarrollar capacidades lingüísticas complejas para poder adquirir la capacidad para atribuir estados proposicionales a otros agentes. En concreto, la hipótesis establece que existe una conexión a largo del desarrollo del niño entre sus capacidades conversacionales y su capacidad para atribuir actitudes proposicionales. Entender a otros como seres mentales comprende evaluar a esos agentes como siendo responsable de ciertos contenidos conectados con patrones particulares de comportamiento.

Existen varios resultados en psicología del desarrollo que confiere a la hipótesis un lecho sólido de evidencia en la que apoyarse. En general, hay un importante número de estudios que sugieren una fuerte conexión entre determinadas habilidades lingüísticas y la exposición a conversaciones, por un lado, y la capacidad de adscribir estados mentales por el otro. En este sentido, he presentado diferentes estudios que hacen la hipótesis conversacional evaluativa preferible a otras versiones de la visión conversacional (Capítulo 6: sección 5). En general, los niños parecen hacer uso de adscripciones en contextos de disputa y conflicto con mayor frecuencia que en otro tipo de contextos conversa-

cionales. En la misma línea, entornos cooperativos como los juegos de simulación o juego conjuntos parecen estar relacionados con las capacidades explicativas y de atribución. Lo que todas estas situaciones comparten es que disparas diferentes respuestas regulativas que facilitan la aparición de estrategias explicativas y justificativas de la acción.

Además de la capacidad de la hipótesis para dar cuenta de estos resultados empíricos, la hipótesis conversacional evaluativa viene motivada por una concepción filosófica particular de la naturaleza de las adscripciones de actitud proposicional (La visión Evaluativa). De acuerdo con esta visión, los usos paradigmáticos de adscripciones en contextos de explicación son evaluaciones. Estos usos de las adscripciones de actitud proposicional evalúan a los agentes atribuidos en base a diferentes grados de responsabilidad, mérito, demérito o significación hacia un determinado contenido. Esto implica comprender las adscripciones como herramientas para justificación, condena o racionalización en situaciones donde nuestras conductas están permitidas o prohibidas por un conjunto de normas que gobiernan nuestras prácticas sociales. Teniendo este análisis filosófico en mente, hemos podido explicar por qué enfrentarse a los contextos conversacionales mencionados conectan nuestras capacidades para mentalizar con la comunicación lingüística.

La visión evaluativa ha sido presentada como una alternativa a la visión recibida en cognición social según la cual las adscripciones de actitud proposicional son descripciones, esto es, representaciones de entidades internas que reflejan la maquinaria psicológica interna con propósitos explicativos y predictivos. Las razones para promover la visión evaluativa sobre el descriptivismo han sido de dos tipos. Por un lado, el descriptivismo se enfrenta a varios problemas (Capítulo 4). Primero, algunos usos de las expresiones que contienen verbos de actitud proposicional (parentéticos y comunicativos) no tienen un análisis

descriptivo automático y su uso revela una función evaluativa. En segundo lugar, el descriptivismo de la psicología popular no puede dar cuenta de un fenómeno particular; a saber, el carácter persistente de cierto tipos de desacuerdos. Finalmente, algunos tipos de usos explicativos de los predicados mentales están asociados a situaciones donde el atribuidor se enfrenta a comportamientos contra-normativos. Estos rasgos de ciertos usos de los verbos de actitud proposicional sólo tienen sentido cuando se consideran l como evaluaciones de los sujetos como alguien con cierto mérito o responsabilidad que racionaliza la conducta anómala.

Por otro lado, hay varios rasgos de las adscripciones de actitud proposicional que nos permiten inclinar la balanza del lado de la visión evaluativa (Capítulo 5). En primer lugar, los verbos de actitud proposicional comparten algunos rasgos con otras expresiones evaluativas como los términos morales. Por ejemplo, su conexión especial con la acción. En segundo lugar, ciertos fenómenos apuntados por algunas investigaciones empíricas no parecen ser coherentes con una interpretación descriptivista de las atribuciones. Por ejemplo, se esperaría que los humanos fueran más diestros en usar actitudes proposicionales en situaciones explicativas que en situaciones de predicción. En conclusión, he argumentado que hay tanto razones filosóficas como evidencia empírica suficiente para considerar la hipótesis conversacional evaluativa una propuesta seria sobre la relación entre lenguaje y psicología popular. Incluso si alguien se resistiera a abrazar dicha hipótesis, espero haber puesto sobre la mesa argumentos los suficientemente persuasivos para convencer de la necesidad de dirigir al foco del análisis filosófico hacia algunos de los supuestos de fondo en el debate científico sobre el rol del lenguaje en la psicología popular. Un gran número de psicólogos del desarrollo, científicos cognitivos y filósofos de la mente han mantenido de manera acrítica una visión particular de la naturaleza de las adscripciones mentales.

Por tanto, incluso aunque haya razones par resistirse a abrazar la visión evaluativa, los teóricos deberían de hacer frente a los argumentos que he expuesto en esta tesis; o al menos, ser conscientes de la necesidad de evaluar críticamente las asunciones filosóficas detrás de los diferentes puntos de vista empíricos. Una conclusión similar debería extraerse con respecto a otros supuestos de fondo, por ejemplo, la idea de que nuestras capacidades sociales son básicamente una cuestión de explicación y predicción.

Naturalmente, estas conclusiones nos permiten articular diferentes ideas que pueden abrir diferente vías para futuras investigaciones. Desde el punto de vista empírico, las preocupaciones metodológicas sobre los diferentes test de creencia falsa expuestas en el capítulo 6, llaman la atención sobre la necesidad de diseñar diferentes vías empíricas que permitan investigar con mayor refinamiento la distinción entre explicación y predicción en contextos sociales. Por supuesto, más trabajo empírico se necesita para validar o falsar la hipótesis principal defendida; en este sentido, sería necesario diseñar diferentes escenarios experimentales que nos permitan decidir entre hipótesis rivales.

Desde el punto de vista conceptual, hay diferentes líneas de investigación que podrían desarrollarse tomando como punto de partida algunas de las ideas expuestas en este trabajo. En primer lugar, uno podría tomar la estrategia crítica del descriptivismo que he desarrollado en esta tesis como punto de partida para reforzar teóricamente las ideas que he atacado. En segundo lugar, si estoy en lo cierto, y el descriptivismo es una idea generalizada en los debates sobre psicología popular, uno podría explotar las consecuencias teóricas del anti-descriptivismo o la visión evaluativa para otros debates y áreas del estudio de la cognición social. Por ejemplo, el debate acerca de los mecanismos que subyacen a nuestras adscripciones. Finalmente, uno podría considerar enfrentarse a las dos tesis presentadas y su conexión. Alguien podría rechazar el descriptivismo

y abrazar la visión evaluativa de las adscripciones, y sin embargo, dudar de la plausibilidad de la hipótesis conversacional evaluativa. De hecho, el argumento de Andrews presentado en el capítulo 6 podría ser interpretado en esta línea. Ciertamente, aunque he intentado dar respuesta a estos argumentos, hay varias posibles réplicas y resultados experimentales que podrían forzar un desarrollo conceptual de la conexión entre la visión evaluativa de las adscripciones y la hipótesis conversacional evaluativa. A la luz de esto, creo que es justo decir que la investigación filosófica presentada en esta tesis justificaría varias futuras líneas de investigación, algunas de las cuales me gustaría explorar en los años siguientes

References

Juan José Acero and Neftalí Villanueva Fernández. Wittgenstein y la intencionalidad de lo mental. *Análisis filosófico*, 32(2):117–154., 2012.

Karin Aijmer. I think: an english modal particle. In Toril Swan and Olaf Westik, editors, *Modality in Germanic Language: Historical and Comparative Perspectives*, pages 1–48. De Gruyter Mouton, 1997.

Kristin Andrews. How to learn from our mistakes. *Philosophical Explorations*, 7(3):247–263, 2004.

Kristin Andrews. Understanding norms without a theory of mind. *Inquiry*, 52 (5):433–448, 2009.

Kristin Andrews. *Do Apes Read Minds?: Toward a New Folk Psychology*. MIT Press, 2012.

Kristin Andrews. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. Routledge, 2014.

Kristin Andrews. The folk psychological spiral: Explanation, regulation, and language. *Southern Journal of Philosophy*, 53:50–67, 2015.

Ian Apperly and Elizabeth J. Robinson. When can children handle referential opacity? evidence for systematic variation in 5-and 6-year-old children's

reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, 85(4):297–311, 2003.

Ian A. Apperly. *Mindreaders: The Cognitive Basis of "Theory of Mind"*. Taylor & Francis, 2011.

Ian A. Apperly. Can theory of mind grow up? mindreading in adults, and its implications for the development and neuroscience of mindreading. In Simon Baron-Cohen, Helen Tager-Flusberg, and Michael C. Lombardo, editors, *Understanding other minds: Perspectives from developmental social neuroscience*, pages 72–92. Oxford University Press, 2013.

Ian A. Apperly and Elizabeth J. Robinson. Children's mental representation of referential relations. *Cognition*, 67(3):287–309, 1998.

Ian A. Apperly, Kevin Riggs, Andrew Simpson, Claudia Chiavarino, and Dana Samson. Is belief reasoning automatic? *Psychological Science*, 17(10):841–844, 2006.

Ian A. Apperly, Elisa Back, Dana Samson, and Lisa France. The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, 106(3):1093–1108, 2008.

Ian A. Apperly, Frances Warren, Benjamin Andrews, Jay Grant, and Sophie Todd. Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child development*, 82(5):1691–1703, 2011.

Janet Wilde Astington and Jodie A. Baird, editors. *Why Language Matters for Theory of Mind*. Oxford University Press, 2005.

Janet Wilde Astington and Jennifer M. Jenkins. A longitudinal study of the

relation between language and theory-of-mind development. *Developmental psychology*, 35(5):1311, 1999.

John L. Austin. *How to Do Things with Words*. Clarendon Press, 1962.

Alfred J. Ayer Ayer. *Language, Truth and Logic*. London: V. Gollancz, 1936.

Renée Baillargeon, Rose M. Scott, and Zijing He. False-belief understanding in infants. *Trends in Cognitive Sciences*, 14(3):110–118, 2010.

Dorit Bar-On. *Speaking My Mind: Expression and Self-Knowledge*. Oxford University Press UK, 2004.

Stephen J. Barker. Is value content a component of conventional implicature? *Analysis*, 60(267):268–279, 2000.

Simmon Baron-Cohen. *Mindblindness an Essay on Autism and "Theory of Mind"*. 1995.

Simmon Baron-Cohen. The evolution of a theory of mind. In Michael C. Corballis and S. E. G. Lea, editors, *The Descent of Mind Psychological Perspectives on Hominid Evolution*. New York Oxford University Press, 1999.

Karen Bartsch. The role of experience in children's developing folk epistemology: review and analysis from the theory–theory perspective. *New Ideas in Psychology*, 20(2):145–161, 2002.

Karen Bartsch and Henry Wellman. Young children's attribution of action to beliefs and desires. *Child development*, pages 946–964, 1989.

Karen Bartsch and Henry Wellman. *Children Talk about the Mind*. Oxford University Press, 1995.

James R. Beebe. A knobe effect for belief ascriptions. *Review of Philosophy and Psychology*, 4(2):235–258, 2013.

James R. Beebe and Wesley Buckwalter. The epistemic side-effect effect. *Mind and Language*, 25(4):474–498, 2010.

Nuel Belnap. Declaratives are not enough. *Philosophical Studies*, 59(1):1–30, 1990.

Jose Luis Bermudez. *Thinking Without Words*. Oxford University Press, 2003a.

Jose Luis Bermudez. The domain of folk psychology. In Anthony O'Hear, editor, *Royal Institute of Philosophy Supplement*, pages 25–48. Cambridge University Press, 2003b.

Jose Luis Bermudez. Arguing for eliminativism. In Brian L. Keeley, editor, *Paul Churchland*. Cambridge University Press, 2005.

Jose Luis Bermudez. Can non-linguistic animals think about thinking? In J. Beck and K. Andrews, editors, *The Routledge Companion to Animal Minds*. Routledge, forthcoming.

Derek Bickerton. *Roots of language*. Language Science Press, 1981/2015.

Paul Bloom. *How children learn the meanings of words*. MIT press Cambridge, MA, 2000.

Christophe Boesch. Cooperative hunting roles among tai chimpanzees. *Human Nature*, 13(1):27–46, 2002.

Paul A. Boghossian. The status of content. *Philosophical Review*, 99(2):157–84, 1990.

Vivian Bohl and Wouter van den Bos. Toward an integrative account of social cognition: marrying theory of mind and interactionism to study the interplay of type 1 and type 2 processes. *Frontiers in human neuroscience*, 6:274, 2012.

Emma Borg. If mirror neurons are the answer, what was the question? *Journal of Consciousness Studies*, 14(8):5–19, 2007.

George Botterill. Folk psychology and theoretical status. In Peter Carruthers and Peter K. Smith, editors, *Theories of Theories of Mind*, pages 105–118. Cambridge University Press, 1996.

Carlos Bouza. Pikara mgazine: "todo lo que escribo es el resultado de una reflexión colectiva", february 2015. URL `http://www.pikaramagazine.com/`.

Jane R. Brown, Nancy Donelan-McCall, and Judy Dunn. Why talk about mental states? the significance of children's conversations with friends, siblings, and mothers. *Child development*, 67(3):836–849, 1996.

Jerome S. Bruner. *Acts of Meaning*. Harvard University Press, 1990.

Filip Buekens. Faultless disagreement, assertions and the affective-expressive dimension of judgments of taste. *Philosophia*, 39(4):637–655, 2011.

David Buttelmann, Malinda Carpenter, and Michael Tomasello. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2):337–342, 2009.

Josep Call and Michael Tomasello. Does the chimpanzee have a theory of mind? 30 years later. *Trends in cognitive sciences*, 12(5):187–192, 2008.

Stephanie M. Carlson and Louis J. Moses. Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4):1032–1053, 2001.

Stephanie M. Carlson, Louis J. Moses, and Hollie R. Hix. The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, 69(3):672–691, 1998.

Jeremy I. M. Carpendale and Charlie Lewis. Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, 27(1):79–96, 2004.

Jeremy I. M. Carpendale and Charlie Lewis. *How children develop social understanding*. Blackwell Publishing, 2006.

Peter Carruthers. *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*. Oxford: Clarendon Press, 2006.

Peter Carruthers. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. OUP Oxford, 2011.

Peter Carruthers. Mindreading in infancy. *Mind and Language*, 28(2):141–172, 2013.

Peter Carruthers and Jill Boucher. *Language and Thought*. Cambridge University Press, 1998.

Robin Carston. *Thoughts and Utterances*. Blackwell, 2002.

Bianca Cepollaro and Isidora Stojanovic. Hybrid evaluatives: In defense of a presuppositional account. *Grazer Philosophische Studien*, 93(3):458–488, 2016.

Nate Charlow. The problem with the frege–geach problem. *Philosophical Studies*, 167(3):635–665, 2014.

Nate Charlow. Prospects for an expressivist theory of meaning. *Philosophers' Imprint*, 15(23):1–43, 2015.

Matthew Chrisman. From epistemic contextualism to epistemic expressivism. *Philosophical Studies*, 135(2):225–254, 2007.

Matthew Chrisman. From epistemic expressivism to epistemic inferentialism. In Adrian Haddock, Alan Millar, and Duncan Pritchard, editors, *Social Epistemology*. Oxford University Press, 2010.

Matthew Chrisman. Epistemic expressivism. *Philosophy Compass*, 7(2):118–126, 2012.

Marvin M. Chun and Yuhong Jiang. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive psychology*, 36 (1):28–71, 1998.

Paul M. Churchland. Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78:67–90, 1981.

Paul M. Churchland. *Matter and Consciousness*. MIT Press, 1984.

Andy Clark. Magic words: How language augments human computation. In P. Carruthers and J. Boucher, editors, *Language and Thought: Interdisciplinary Themes*, pages 162–183. Cambridge, 1998.

Matthew Van Cleave and Christopher Gauker. Linguistic practice and false-belief tasks. *Mind and Language*, 25(3):298–328, 2010.

Russell W. Clement and Joachim Krueger. Social categorization moderates social projection. *Journal of experimental social psychology*, 38(3):219–231, 5 2002.

Stewart Cohen. How to be a fallibilist. *Philosophical Perspectives*, 2(n/a):91–123, 1988.

Stewart Cohen. Contextualism, skepticism, and the structure of reasons. *Philosophical Perspectives*, 13(s13):57–89, 1999.

Daniel Cohnitz and Teresa Marques. Disagreements. *Erkenntnis*, 79(1):1–10, 2014.

Benjamin A Converse, Shuhong Lin, Boaz Keysar, and Nicholas Epley. In the mood to get over yourself: mood affects theory-of-mind use. *Emotion*, 8(5): 725, 2008.

David Copp. Realist-expressivism and conventional implicature. *Oxford Studies in Metaethics*, 4:167–202, 2009.

Alexandra L. Cutting and Judy Dunn. Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, 70(4):853–865, 1999.

Donald Davidson. Mental events. In L. Foster and J. W. Swanson, editors, *Essays on Actions and Events*, pages 207–224. Clarendon Press, 1970.

Donald Davidson. Thought and talk. In Samuel D. Guttenplan, editor, *Mind and Language*. Clarendon Press, 1975.

Donald Davidson. Rational animals. *Dialectica*, 36(4):317–28, 1982.

Leon C. de Bruin and Derek W. Strijbos. Folk psychology without principles: An alternative to the belief-desire model of action interpretation. *Philosophical Explorations*, 13(3):257–274, 2010.

Hanne De Jaegher and Ezequiel Di Paolo. Participatory sense-making. *Phenomenology and the Cognitive Sciences*, 6(4):485–507, 2007.

Dan López de Sa. Expressing disagreement: A presuppositional indexical contextualist relativist account. *Erkenntnis*, 80(1):153–165, 2015.

Jill G. de Villiers. Iq can language acquisition give children a point ofview. *Why language matters for theory of mind*, 186, 2005.

Jill G. De Villiers. The interface of language and theory of mind. *Lingua. International review of general linguistics. Revue internationale de linguistique generale*, 117(11):1858–1878, 2007.

Jill G. De Villiers and Peter A. De Villiers. Linguistic determinism and the understanding of false. *Children's reasoning and the mind*, pages 191–228, 2000.

Jill G. de Villiers and Peter A. de Villiers. Language for thought: Coming to understand false beliefs. *Language in mind: Advances in the study of language and thought*, pages 335–384, 2003.

Jill G. De Villiers and Jennie E. Pyers. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development*, 17(1):1037–1060, 2002.

Susanne A Denham, Daniel Zoller, and Elizabeth A Couchoud. Socialization of preschoolers' emotion understanding. *Developmental psychology*, 30(6):928, 1994.

Daniel C. Dennett. *Brainstorms: philosophical essays on mind and psychology*. Bradford Books, 1978a.

Daniel C. Dennett. Beliefs about beliefs (commentary on premack, et al.). *Behavioral and Brain Sciences*, 1(4):568, 1978b.

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

Daniel C Dennett. Real patterns. *The journal of Philosophy*, 88(1):27–51, 1991b.

Keith DeRose. Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, 52(4):913–929, 1992.

Keith DeRose. Solving the skeptical problem. *Philosophical Review*, 104(1): 1–52, 1995.

Miguel dos Santos, Daniel J. Rankin, and Claus Wedekind. The evolution of punishment through reputation. *Proceedings of the Royal Society of London B: Biological Sciences*, page rspb20101275, 2010.

Robin Dunbar. On the origin of the human mind. In P. Carruthers and A. Chamberlain, editors, *Evolution and the Human Mind*, pages 238–253. Cambridge University Press, 2000.

Robin Dunbar. The social brain : Mind, language, and society in evolutionary perspective. *Annual Review of Anthropology*, 32(1):163–181, 2003.

Judy Dunn. Changing minds and changing relationships. In Charlie Lewis and Peter Mitchell, editors, *Children's early understanding of mind: Origins and development*, pages 297–310. Psychology Press, 1994.

Judy Dunn and Jane R. Brown. Early conversations about causality: Content, pragmatics and developmental change. *British Journal of Developmental Psychology*, 11(2):107–123, 1993.

Judy Dunn and Penny Munn. Development of justification in disputes with mother and sibling. *Developmental Psychology*, 23(6):791, 1987.

Judy Dunn, Jane R. Brown, Cheryl Slomkowski, Caroline Tesla, and Lise Youngblade. Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child development*, 62(6):1352–1366, 1991.

Carol S. Dweck, Chi-yue Chiu, and Ying-yi Hong. Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological inquiry*, 6(4):267–285, 1995.

Ernst Fehr and Simon Gächter. Altruistic punishment in humans. *Nature*, 415 (6868):137–40, 01 2002.

Adam Feltz. The knobe effect: A brief overview. *Journal of Mind and Behavior*, 28(3-4):265–277, 2007.

Víctor Fernández-Castro. Dar forma a la mente. *THÉMATA. Revista de Filosofía*, (51):433–442, 2015a.

Víctor Fernández-Castro. The normative dimension of folk psychology. In J. Díez, M. García-Carpintero, J. Martínez, and S. Oms, editors, *VIII Conference of the Spanish Society for Logic, Methodology and Philosophy of Science*. University of Barcelona, 2015b.

Víctor Fernández-Castro. Inner speech in action. *Pragmatics and Cognition*, 23 (2), 2017.

Hartry Field. Epistemology without metaphysics. *Philosophical Studies*, 143(2): 249–290, 2009.

Stephen Finlay. The conversational practicality of value judgement. *Journal of Ethics*, 8(3):205–223, 2004.

Guy Fletcher. Hybrid views in metaethics: Pragmatic views. *Philosophy Compass*, 9(12):848–863, 2014.

Jerry A. Fodor. *The Language of Thought*. Harvard University Press, 1975.

Jerry A. Fodor. Propositional attitudes. *The Monist*, 61(October):501–23, 1978.

Jerry A. Fodor. Fodor's guide to mental representation: The intelligent auntie's vade-mecum. *Mind*, 94(373):76–100, 1985.

Jerry A. Fodor. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press, 1987.

Jerry A. Fodor. A theory of the child's theory of mind. *Cognition*, 44(3):283–296, 1992.

Jerry A. Fodor. *Concepts: Where cognitive science went wrong*. Oxford University Press, 1998.

Rachel C. Foote and Heather A. Holmes-Lonergan. Sibling conflict and theory of mind. *British Journal of Developmental Psychology*, 21(1):45–58, 2003.

María José Frápolli. *The Nature of Truth*. Springer, 2013.

María José Frápolli and Neftalí Villanueva Fernández. Minimal expressivism. *Dialectica*, 66(4):471–487, 2012.

Norman H. Freeman and Hazel Lacohée. Making explicit 3-year-olds' implicit competence with their own false beliefs. *Cognition*, 56(1):31–60, 7 1995.

David Furrow, Chris Moore, Jane Davidge, and Lorraine Chiasson. Mental terms in mothers' and children's speech: Similarities and relationships. *Journal of child Language*, 19(03):617–631, 1992.

Shaun Gallagher. The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies*, 8(5-7):83–108, 2001.

Shaun Gallagher. Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2):535–543, 2008.

Shaun Gallagher and Daniel D. Hutto. Understanding others through primary interaction and narrative practice. In J. Zlatev, T. Racine, C. Sinha, and E. Itkonen, editors, *The Shared Mind: Perspectives on Intersubjectivity.*, pages 17–38. John Benjamins, 2008.

Pamela W. Garner, Diane Carlson Jones, Gaylyn Gaddy, and Kimberly M Rennie. Low-income mothers' conversations about emotions and their children's emotional competence. *Social Development*, 6(1):37–52, 1997.

Christopher Gauker. *Thinking Out Loud: An Essay on the Relation Between Thought and Language*. Princeton University Press, 1995.

Christopher Gauker. *Words Without Meaning*. MIT Press, 2003.

Dedre Gentner and Susa Goldin-Meadow, editors. *Language in mind: Advances in the study of language and thought*. MIT Press, 2003.

Tim P. German and Jessica A. Hehman. Representational and executive selection resources in 'theory of mind': Evidence from compromised belief-desire reasoning in old age. *Cognition*, 101(1):129–152, 2006.

Allan Gibbard. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press, 1990.

Allan Gibbard. *Thinking How to Live*. Harvard University Press, 2003.

Allan Gibbard. *Meaning and Normativity*. Oxford University Press Uk, 2012.

Cliff Goddard. Thinking across languages and cultures: Six dimensions of variation. *Cognitive Linguistic*, 14(2/3):109–140, 2003.

Peter Goldie. There are reasons and reasons. In Daniel D. Hutto and Matthew Ratcliffe, editors, *Folk Psychology Re-Assessed*, pages 103–114. Kluwer/Springer Press, 2007.

Alvin I. Goldman. Interpretation psychologized. *Mind and Language*, 4(3): 161–85, 1989.

Alvin I. Goldman. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford University Press, 2006.

Alvin I. Goldman. Theory of mind. In E. Margolis, R. Samuels, and S. P. Stich, editors, *Oxford Handbook of Philosophy of Cognitive Science*. Oxford University Press, 2012.

Susan Golombok and Robyn Fivush. *Gender Development*. Cambridge University Press, 1994.

Juan Carlos Gomez. *Apes, Monkeys, Children, and the Growth of Mind*. Harvard University Press, 2004.

Antoni Gomila. *Verbal minds : language and the architecture of cognition*. Elsevier, 2012.

Alison Gopnik. Theories and modules; creation myths, developmental realities, and neurath's boat. In P. Carruthers and P. K. Smith, editors, *Theories of Theories of Mind*, page 169. Cambridge University Press, 1996.

Alison Gopnik. The theory theory as an alternative to the innateness hypothesis. In Louise M. Antony, editor, *Chomsky and His Critics*, pages 238–254. Blackwell, 2003.

Alison Gopnik and Janet W. Astington. Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59(1):26–37, 1988.

Alison Gopnik and Andrew Meltzoff. *Words, Thoughts, and Theories*. MIT Press, 1997.

Alison Gopnik and Henry M. Wellman. The theory theory. In L. A. Hirschfeld and S. A. Gelman, editors, *Mapping the mind: Domain specificity in cognition and culture*, pages 257–293. Cambridge University Press, New York, NY, US, 1994.

Robert M. Gordon. Sympathy, simulation, and the impartial spectator. *Ethics*, 105(4):727–742, 1995.

Robert M. Gordon. Simulation and the explanation of action. In K. R. Stueber and H. H. Kogaler, editors, *Empathy and Agency: The Problem of Understanding in the Human Sciences*. Boulder: Westview Press, 2000.

Anthony G. Greenwald, Colin Tucker Smith, N. Sriram, Yoav Bar-Anan, and Brian A. Nosek. Implicit race attitudes predicted vote in the 2008 u.s. presidential election. *Analyses of Social Issues and Public Policy*, 9(1):241–253, 2009.

Paul Grice. *Logic and Conversation*, pages 41–58. Harvard University Press, 1967.

Courtney Melinda Hale and Helen Tager-Flusberg. The influence of language on theory of mind: A training study. *Developmental science*, 6(3):346–359, 2003.

Francesca G. E. Happé. The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66(3):843–855, 1995.

Richard M. Hare. *The Language of Morals*. Oxford Clarendon Press, 1952.

Gilbert Harman. Studying the chimpanzee's theory of mind (commentary on premack, et al.). *Behavioral and Brain Sciences*, 1(4):576, 1978.

Paul L. Harris. Desires, beliefs, and language. In P. Carruthers and P. K. Smith, editors, *Theories of Theories of Mind*, pages 200–220. Cambridge University Press, 1996.

Paul L. Harris. Conversation, pretense, and theory of mind. In J. W. Astington and J. Baird, editors, *Why Language Matters for Theory of Mind*, pages 70–83. Oxford University Press, New York, 2005.

Paul L Harris, Marc de Rosnay, and Francisco Pons. Language and children's understanding of mental states. *Current directions in psychological science*, 14(2):69–73, 2005.

Sally Haslanger. Distinguished lecture: Social structure, narrative and explanation. *Canadian Journal of Philosophy*, 45(1):1–15, 2015.

Max Hayward. How to be an expressivist about practical rationality. Presented in New Directions for Expressivism, Sheffield, 2016.

Jane Heal. Simulation, theory, and content. In Peter Carruthers and Peter K. Smith, editors, *Theories of Theories of Mind*, pages 75–89. Cambridge University Press, 1996.

Jane Heal. Co-cognition and off-line simulation: Two ways of understanding the simulation approach. *Mind and Language*, 13(4):477–498, 1998.

Fritz Heider. *The Psychology of Interpersonal Relations*. Lawrence Erlbaum Associates, 1958.

Joseph Henrich. Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, 53(1):3–35, 2004.

Joseph Henrich, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer, and John Ziker. Costly punishment across human societies. *Science*, 312 (5781):1767, 2006.

Cecilia Heyes. False belief in infancy: a fresh look. *Developmental Science*, 17 (5):647–659, 2014.

Cecilia Heyes. Apes submentalise. *Trends in Cognitive Sciences*, 21(1):1–2, 2017.

Terence Horgan and Mark Timmons. Troubles for new wave moral semantics: The 'open question argument' revived. *Philosophical Papers*, 21(3):153–175, 1992.

Claire Hughes and Judy Dunn. "pretend you didn't know": Preschoolers' talk about mental states in pretend play. *Cognitive Development*, 12(4):477–497, 1997.

Claire Hughes and Judy Dunn. Understanding mind and emotion: longitudinal associations with mental-state talk between young friends. *Developmental psychology*, 34(5):1026, 1998.

Claire Hughes, Keiko K. Fujisawa, Rosie Ensor, Serena Lecce, and Rachel Marfleet. Cooperation and conversations about the mind: A study of individual differences in 2-year-olds and their siblings. *British Journal of Developmental Psychology*, 24(1):53–72, 2006.

Daniel Hutto. Fictionalism about folk psychology. *The Monist*, 96(4):582–604, 2013.

Daniel D. Hutto. The limits of spectatorial folk psychology. *Mind and Language*, 19(5):548–73, 2004.

Daniel D. Hutto. *Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons*. A Bradford Book, 2008a.

Daniel D. Hutto and Matthew Ratcliffe. *Folk Psychology Re-Assessed*. Kluwer/Springer Press, 2007.

Ray S. Jackendoff. How language helps us think. *Pragmatics and Cognition*, 4 (1):1–34, 1996.

Jennifer M. Jenkins and Janet Wilde Astington. Theory of mind and social behavior: Causal models tested in a longitudinal study. *Merrill-Palmer Quarterly (1982-)*, pages 203–220, 2000.

M. Jennifer Jenkins and Janet Wilde Astington. Cognitive factors and family structure associated with theory of mind development in young children. *Developmental psychology*, 32(1):70, 1996.

E.E. Jones and R.E. Nisbett. *The actor and the observer: divergent perceptions of the causes of behavior*. General Learning Press, 1971.

Marta Jorba and Agustin Vicente. Cognitive phenomenology, access to contents, and inner speech. *Journal of Consciousness Studies*, 21(9-10):74–99, 2014.

Charles W. Kalish and Christopher A. Lawson. Development of social category representations: Early appreciation of roles and deontic relations. *Child Development*, 79(3):577–593, 2008.

Deepthi Kamawar and David R Olson. Children's understanding of referentially opaque contexts: The role of metarepresentational and metalinguistic ability. *Journal of Cognition and Development*, 10(4):285–305, 2009.

Deepthi Kamawar and David R Olson. Thinking about representations: The case of opaque contexts. *Journal of experimental child psychology*, 108(4): 734–746, 2011.

Klemens Kappel and Emil F. L. Moeller. Epistemic expressivism and the argument from motivation. *Synthese*, 191(7):1–19, 2013.

Lauri Karttunen. Presuppositions of compound sentences. *Linguistic inquiry*, 4 (2):169–193, 1973.

Boaz Keysar, Shuhong Lin, and Dale J. Barr. Limits on theory of mind use in adults. *Cognition*, 89(1):25–41, 2003.

Joshua Knobe. Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, 16(2):309–325, 2003.

Joshua Knobe. Person as scientist, person as moralist. *Behavioral and Brain Sciences*, 33(04):315–329, 2010.

Birgit Knudsen and Ulf Liszkowski. Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science*, 15(1):113–122, 2012.

Joanna Korman and Bertram F. Malle. Grasping for traits or reasons? how people grapple with puzzling social behaviors. *Personality and Social Psychology Bulletin*, 42(11):1451–1465, 2016.

Carla Krachun, Malinda Carpenter, Josep Call, and Michael Tomasello. A competitive nonverbal false belief task for children and apes. *Developmental Science*, 12(4):521–535, 2009.

Christopher Krupenye, Fumihiro Kano, Satoshi Hirata, Josep Call, and Michael Tomasello. Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308):110–114, 2016.

Christopher E. Lalonde and Michael J. Chandler. Children's understanding of interpretation. *New Ideas in Psychology*, 20(2):163–198, 2002.

Mark N. Lance and John O'Leary-Hawthorne. *The Grammar of Meaning*. Cambridge University Press, 1997.

Jane Lavelle. Cross cultural considerations in social cognition. In J. Kiverstein, editor, *The Routledge Handbook to The Social Mind*. Routledge, 2016.

Takie Sugiyama Lebra. Culture, self, and communication in japan and the united states. *Communication in Japan and the United States*, 51:87, 1993.

Cristine H. Legare, Susan A. Gelman, and Henry M. Wellman. Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child development*, 81(3):929–944, 2010.

Alan M. Leslie. Pretense and representation: The origins of "theory of mind.". *Psychological Review*, 94(4):412–426, 1987.

Alan M. Leslie. Pretending and believing: Issues in the theory of tomm. *Cognition*, 50(1-3):211–238, 1994.

Alan M. Leslie. How to acquire a 'representational theory of mind'. In D. Sperber, editor, *Metarepresentations*, pages 197–223. Oxford University Press, 2000.

Alan M. Leslie and Pamela Polizzi. Inhibitory processing in the false belief task: Two conjectures. *Developmental Science*, 1(2):247–253, 1998.

Alan M. Leslie and Daniel Roth. What autism teaches us about metarepresentation. 1993.

Alan M. Leslie and Laila Thaiss. Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, 43(3):225–251, 1992.

Ivan Leudar and Alan Costall, editors. *Theory of Mind : The Madness Behind the Method*. Palgrave Macmillan, 2009.

Ivan Leudar, Alan Costall, and Dave Francis. Theory of mind: a critical assessment. *Theory and Psychology*, 14(5):571–578, 10 2004.

Charlie Lewis, Norman H Freeman, Chrystalla Kyriakidou, Katerina Maridaki-Kassotaki, and Damon M Berridge. Social influences on false belief access:

specific sibling influences or general apprenticeship? *Child development*, 67 (6):2930–2947, 1996.

David Lewis. Attitudes de dicto and de se. *Philosophical Review*, 88(4):513–543, 1979.

David Liu, Henry M. Wellman, Twila Tardif, and Mark A. Sabbagh. Theory of mind development in chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Developmental psychology*, 44(2): 523, 2008.

Heidemarie Lohmann and Michael Tomasello. The role of language in the development of false belief understanding: A training study. *Child development*, 74(4):1130–1144, 2003.

Charles G. Lord, Lee Ross, and Mark R. Lepper. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology*, 37(11):2098, 1979.

Jason Low, William Drummond, Andrew Walmsley, and Bo Wang. Representing how rabbits quack and competitors act: limits on preschoolers' efficient ability to track perspective. *Child development*, 85(4):1519–1534, 2014.

John MacFarlane. *Assessment Sensitivity: Relative Truth and its Applications*. Oxford University Press, 2014.

John L. Mackie. *Ethics: Inventing Right and Wrong*. Penguin Books, 1977.

Heidi L. Maibom. Social systems. *Philosophical Psychology*, 20(5):557–578, 2007.

Bertram F. Malle. How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review*, 3(1):23–48, 1999.

Bertram F. Malle. *How the Mind Explains Behavior: Folk Explanations, Meaning, and Social Interaction*. 2004.

Bertram F. Malle. Attribution theories: How people make sense of behavior. *Theories in Social Psychology.*, pages 72–95, 2011.

Bertram F. Malle, Joshua Knobe, and S. Nelson. Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology*, 93:491–514, 2007.

Matteo Mameli. Mindreading, mindshaping, and evolution. *Biology and Philosophy*, 16(5):595–626, 2001.

Fernando Martínez Manrique and Agustín Vicente. Hablar para pensar: sobre el uso del lenguaje en el pensamiento. *Análisis filosófico*, 28(1):91–112, 2008.

Fernando Martínez Manrique and Agustín Vicente. What the...!'the role of inner speech in conscious thought. *Journal of Consciousness Studies*, 17(9-1):141–167, 2010.

Fernando Martínez Manrique and Agustín Vicente. The activity view of inner speech. *Frontiers in psychology*, 6:232, 2015.

Gregory McCulloch. Dennett's little grains of salt. *Philosophical Quarterly*, 40 (158):1–12, 1990.

Victoria McGeer. Psycho-practice, psycho-theory and the contrastive case of autism. how practices of mind become second-nature. *Journal of Consciousness studies*, 8(5-6):109–132, 2001.

Victoria McGeer. The regulative dimension of folk psychology. In Daniel D. Hutto and Matthew Ratcliffe, editors, *Folk Psychology Re-Assessed*, pages 137–156. Kluwer/Springer Press, 2007.

Victoria McGeer. Mind-making practices: The social infrastructure of self-knowing agency and responsibility. *Philosophical Explorations*, 18(2):259–281, 2015.

Elizabeth Meins, Charles Fernyhough, Fiona Johnson, and Jane Lidstone. Mind-mindedness in children: Individual differences in internal-state talk in middle childhood. *British Journal of Developmental Psychology*, 24(1):181–196, 2006.

Hugo Mercier and Dan Sperber. Why do humans reason? arguments for an argumentative theory. *Behavioral and brain sciences*, 34(02):57–74, 2011.

Joan G. Miller. Culture and the development of everyday social explanation. *Journal of personality and social psychology*, 46(5):961–978, 1984.

Ruth G. Millikan. *Varieties of Meaning: The 2002 Jean Nicod Lectures*. MIT Press, 2004.

John C. Mitani, David P. Watts, and Martin N. Muller. Recent developments in the study of wild chimpanzee behavior. *Evolutionary Anthropology: Issues, news, and reviews*, 11(1):9–25, 2002.

Steven Mithen. Mind, brain and material culture: An archaeological perspective. In Carruthers and A. Chamberlain, editors, *Evolution and the Human Mind*, pages 207–217. Cambridge University Press, 2000.

Chris Moore, David Furrow, Lorraine Chiasson, and Maria Patriquin. Developmental relationships between production and comprehension of mental terms. *First language*, 14(40):001–17, 1994.

Michael W. Morris and Kaiping Peng. Culture and cause: American and chinese attributions for social and physical events. *Journal of Personality and Social psychology*, 67(6):949, 1994.

Adam Morton. Folk psychology is not a predictive device. *Mind*, 105(417): 119–37, 1996.

Adam Morton. *The Importance of Being Understood: Folk Psychology as Ethics*. Routledge, 2002.

Adam Morton. Folk psychology. In B. McLaughlin, A. Beckermann, and S. Walter, editors, *The Oxford Handbook of Philosophy of Mind*. OUP Oxford, 2009.

Mika Naito and Kayo Koyama. The development of false-belief understanding in japanese children: Delay and difference? *International Journal of Behavioral Development*, 30(4):290–304, 2006.

Shaun Nichols and Stephen P. Stich. *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press, 2003.

Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175, 1998.

R. Nisbett. *The Geography of Thought: How Asians and Westerners Think Differently...and Why*. Free Press, 2004.

Andy Norman. Why we reason: intention-alignment and the genesis of human rationality. *Biology & Philosophy*, 31(5):685–704, 2016. ISSN 1572-8404. doi: 10.1007/s10539-016-9532-4. URL `http://dx.doi.org/10.1007/s10539-016-9532-4`.

Martin A. Nowak and Karl Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685):573–577, 1998.

Sanjida O'Connell and Robin Dunbar. A test for comprehension of false belief in chimpanzees. *Evolution and Cognition*, 9(2):131–140, 2003.

Christopher Y. Olivola and Alexander Todorov. Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, 34(2):83–110, 2010.

David R. Olson. On the origins of beliefs and other intentional states in children. In J. W. Astington, P. Harris, and D. R. Olson, editors, *Developing theories of mind*, pages 414–426. New York: Cambridge University Press, 1988.

Kristine H Onishi and Renée Baillargeon. Do 15-month-old infants understand false beliefs? *Science*, 308(5719):255–258, 2005.

Gloria Origgi and Dan Sperber. Evolution, communication and the proper function of language. In P. Carruthers and A. Chamberlain, editors, *Evolution and the Human Mind*, pages 140–169. Cambridge University Press, 2000.

Eduardo Pérez-Navarro, Víctor Fernández-Castro, Javier González de Prado, and Manuel Heras-Escribano. Not expressivist enough: Normative disagreement about belief attribution. *manuscript*, MS.

Josef Perner. *Understanding the Representational Mind*. Cambridge: MIT Press, 1991.

Josef Perner and Johannes Roessler. From infants' to children's appreciation of belief. *Trends in cognitive sciences*, 16(10):519–525, 2012.

Josef Perner, Susan R. Leekam, and Heinz Wimmer. Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5(2):125–137, 1987.

Josef Perner, Ted Ruffman, and Susan R Leekam. Theory of mind is contagious: You catch it from your sibs. *Child development*, 65(4):1228–1238, 1994.

Josef Perner, Manuel Sprung, Petra Zauner, and Hubert Haider. Want that is understood well before say that, think that, and false belief: A test of de

villiers's linguistic determinism on german–speaking children. *Child development*, 74(1):179–188, 2003.

Josef Perner, Petra Zauner, and Manuel Sprung. What does' that'have to do with point of view? conflicting desires and 'want'in. In J. W. Astington and J. Baird, editors, *Why Language Matters for Theory of Mind*, chapter 11, pages 220–244. Oxford University Press, 2005.

Susan Perry, Mary Baker, Linda Fedigan, Julie GrosLouis, Katherine Jack, KatherineC MacKinnon, JosephH Manson, Melissa Panger, Kendra Pyle, Lisa Rose, et al. Social conventions in wild white-faced capuchin monkeys: Evidence for traditions in a neotropical primate 1. *Current Anthropology*, 44(2): 241–268, 2003.

Candida C Peterson and Michael Siegal. Insights into theory of mind from deafness and autism. *Mind & Language*, 15(1):123–145, 2000.

Candida C. Peterson and Virginia P. Slaughter. Telling the story of theory of mind: Deaf and hearing children's narratives and mental state understanding. *British Journal of Developmental Psychology*, 24(1):151–179, 2006.

Christopher Potts. *The Logic of Conventional Implicatures*. Oxford University Press UK, 2005.

Daniel J. Povinelli and Jennifer Vonk. Chimpanzee minds: Suspiciously human? *Trends in cognitive sciences*, 7(4):157–160, 2003.

Daniel J. Povinelli and Jennifer Vonk. We don't need a microscope to explore the chimpanzee's mind. *Mind and Language*, 19(1):1–28, 2004.

Arthur N. Prior. *Objects of Thought*. Clarendon Press, 1971.

Willard V. Quine. *Word and Object*. MIT Press, 1960.

Hannes Rakoczy. In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese*, pages 1–19, 2015.

Hannes Rakoczy, Delia Bergfeld, Ina Schwarz, and Ella Fizke. Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child development*, 86(2):486–502, 2015.

Frank P. Ramsey. The nature of truth. In N. Rescher and U. Majer, editors, *On Truth: Original Manuscript Materials (1927–1929) from the Ramsey Collection at the University of Pittsburgh*. Springer Netherlands, 1927/1991.

Matthew Ratcliffe. *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*. Palgrave Macmillan, 2007.

Marjorie Rhodes and Amanda C. Brandone. Three-year-olds' theories of mind in actions and words. *Frontiers in psychology*, 5:263, 2014.

Peter J. Richerson and Robert Boyd. *Not By Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press, 2008.

Joel Robbins and Alan Rumsey. Introduction: Cultural and linguistic anthropology and the opacity of other minds. *Anthropological Quarterly*, 81(2): 407–420, 2008.

E. J. Robinson and P. Mitchell. Masking of children's early understanding of the representational mind: Backwards explanation versus prediction. *Child Development*, 66(4):1022–1039, 1995.

Johannes Roessler and Josef Perner. Teleology: belief as perspective. *Understanding Other Minds: Perspectives from developmental social neuroscience*, page 35, 2013.

Michelle Z. Rosaldo. *Knowledge and Passion:*. Cambridge University Press, 1980.

Lee Ross, Mark R. Lepper, and Michael Hubbard. Perseverance in self-perception and social perception: biased attributional processes in the debriefing paradigm. *Journal of personality and social psychology*, 32(5):880, 1975.

Paula Rubio-Fernández and Bart Geurts. How to pass the false-belief task before your fourth birthday. *Psychological Science*, 24(1):27–33, 2013.

Ted Ruffman and Josef Perner. Do infants really understand false belief? *Trends in Cognitive Sciences*, 9(10):462–463, 2005.

Ted Ruffman, Josef Perner, Mika Naito, Lindsay Parkin, and Wendy A Clements. Older (but not younger) siblings facilitate false belief understanding. *Developmental psychology*, 34(1):161, 1998.

Ted Ruffman, Lance Slade, and Elena Crowe. The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child development*, 73(3):734–751, 2002.

James Russell. "can we say...?" children's understanding of intensionality. *Cognition*, 25(3):289–308, 1987.

Gilbert Ryle. *The Concept of Mind. Hutchinson's University*. 1949.

Gilbert Ryle. *On Thinking*. Blackwell, 1979.

Joanne Scheibman. Local patterns of subjectivity in person and verb type in american english coversation. In J. L. Bybee and P. J. Hopper, editors, *Local patterns of subjectivity in person and verb type in American English conversation*, pages 62–89. John Benjamins, 2001.

B. B. Schieffelin. Speaking only your own mind: Reflections on talk, gossip and intentionality in bosavi. *Anthropological Quarterly*, 81(2):431–441, 2008.

Brian J. Scholl and Alan M. Leslie. Modularity, development and "theory of mind". *Mind and Language*, 14(1):131–153, 1999.

Gabriel Segal. Representing representations. In Peter Carruthers and Jill Boucher, editors, *Language and Thought*, pages 146–161. Cambridge University Press, 1998.

Wilfrid S. Sellars. *Empiricism and the Philosophy of Mind*. Harvard University Press, 1956.

Karen Shanton and Alvin Goldman. Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(4):527–538, 2010.

Anne-Marie Simon-Vandenbergen. I think and its dutch equivalents in parliamentary debates. In J. Stig and S. Oksefjell, editors, *Corpora and cross-linguistic research: Theory, method and case studies*, pages 297–317. Atlanta: Rodopi, 1998.

Cheryl L. Slomkowski and Judy Dunn. Arguments and relationships within the family: Differences in young children's disputes with mother and sibling. *Developmental Psychology*, 28(5):919, 1992.

Marc Slors. Folk-psychology as reconstruction. forthcoming.

Michael Smith. *The Moral Problem*. Blackwell, 1994.

Peter K. Smith. Language and the evolution of mind-reading. In P. Carruthers and P. K. Smith, editors, *Theories of Theories of Mind*, page 344. Cambridge University Press, 1996.

Victoria Southgate, Atsushi Senju, and Gergely Csibra. Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18 (7):587–592, 2007.

Shannon Spaulding. Do you see what i see? how social differences influence social interpretation. Venue.

Dan Sperber. *Metarepresentations*. Oxford University Press, 2000.

Dan Sperber and Deidre Wilson. *Relevance: Communication and Cognition*. Wiley, 1996.

Manuel Sprung, Josef Perner, and Peter Mitchell. Opacity and discourse referents: Object identity and object properties. *Mind and Language*, 22(3): 215–245, 2007.

Robert Stalnaker. *Context*. Oxford University Press, 2014.

Jason Stanley. *How Propaganda Works*. Princeton University Press, 2015.

Charles Leslie Stevenson. *Ethics and Language*. London: Oxford University PRess, 1944.

Caj Strandberg. A dual aspect account of moral language. *Philosophy and Phenomenological Research*, 84(1):87–122, 2012.

Derek W. Strijbos and Leon C. de Bruin. Making folk psychology explicit. *Philosophia*, 40(1):139–163, 2012a.

Derek W. Strijbos and Leon C. de Bruin. Making folk psychology explicit: The relevance of robert brandom's philosophy for the debate on social cognition. *The Philosopher's IndexPhilosophia: Philosophical Quarterly of Israel*, 40(1), 2012b.

Derek W. Strijbos and Leon C. de Bruin. Reason attribution without belief-desire ascription. *Grazer Philosophische Studien*, 86(1):157–180, 2013a.

Derek W. Strijbos and Leon C. de Bruin. Universal belief-desire psychology? a dilemma for theory theory and simulation theory. *Philosophical Psychology*, 26(5):744–764, 2013b.

Andrew Surtees, Stephen A. Butterfill, and Ian A. Apperly. Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, 30(1):75–86, 2012.

David Tannenbaum, Peter H Ditto, and David A Pizarro. Different moral values produce different judgments of intentional action. *Unpublished manuscript, University of California-Irvine*, 2007.

Julia Tanney. *Rules, Reason and Self-Knowledge*. Harvard University Press, London, 2013.

Twila Tardif and Henry M Wellman. Acquisition of mental state language in mandarin-and cantonese-speaking children. *Developmental Psychology*, 36(1): 25, 2000.

Caroline Tesla and Judy Dunn. Getting along or getting your own way: The development of young children's use of argument in conflicts with mother and sibling. *Social Development*, 1(2):107–121, 1992.

Sandra A. Thompson and Anthony Mulac. A quantitative perspective on the grammaticization of epistemic parentheticals in english. *Approaches to grammaticalization*, 2:313–329, 1991.

Uku Tooming. Beliefs and desires: from attribution to evaluation. *Philosophia*, pages 1–11, 2016.

Uku Tooming. Mental state attribution for interactionism. *Studia Philosophica Estonica*, forthcoming.

Alain J.-P. C. Tschudin. Belief attribution tasks with dolphins: What social minds can reveal about animal rationality. In Susan Hurley and Matthew Nudds, editors, *Rational Animals?* Oxford University Press, 2006.

James O. Urmson. Parenthetical verbs. *Mind*, 61(244):480–496, 1952.

Kevin Uttich and Tania Lombrozo. Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition*, 116(1):87–100, 2010.

Carel P. Van Schaik, Marc Ancrenaz, Gwendolyn Borgen, Birute Galdikas, Cheryl D Knott, Ian Singleton, Akira Suzuki, Sri Suci Utami, and Michelle Merrill. Orangutan cultures and the evolution of material culture. *Science*, 299(5603):102–105, 2003.

Neftalí Villanueva Fernández. Wittgenstein: descripciones y estados mentales. In J. J. Acero, editor, *Guia Comares de Wittgenstein*. Comares, 2017a.

Neftalí Villanueva Fernández. Expressivismo y semántica. In D. Pérez Chico, editor, *Perspectivas en la filosofía del lenguaje vol II*. Prensas Universitarias de Zaragoza, 2017b.

Penelope G. Vinden. Junin quechua children's understanding of mind. *Child Development*, 67(4):1707–1716, 1996.

Penelope G. Vinden. Understanding minds and evidence for belief: A study of mofu children in cameroon. *International Journal of Behavioral Development*, 26(5):445–452, 2002.

Barbara Von Eckardt. Folk psychology. *A companion to the philosophy of mind. Blackwell, Cambridge*, pages 300–307, 1994.

Peter C. Wason. Reasoning. In *New Horizons in Psychology*, pages 135–151. Penguin Books, 1966.

Anne C. Watson, Kathleen M. Painter, and Marc H. Bornstein. Longitudinal relations between 2-year-olds' language and 4-year-olds' theory of mind. *Journal of Cognition and Development*, 2(4):449–457, 11 2001.

Claus Wedekind and Manfred Milinski. Cooperation through image scoring in humans. *Science*, 288(5467):850–852, 2000.

Henry M Wellman. *Making minds: How theory of mind develops*. Oxford University Press, 2014.

Henry M. Wellman, David Cross, and Julanne Watson. Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, 72(3): 655–84, 2001.

H.M. Wellman. *The Child's Theory of Mind*. A Bradford book. MIT Press, 1990. ISBN 9780262730990. URL `https://books.google.es/books?id=O5loPwAACAAJ`.

Evan Westra. Spontaneous mindreading: A problem for the two-systems account. *Synthese*, pages 1–23, forthcoming.

Evan Westra and Peter Carruthers. Pragmatic development explains the theory-of-mind scale. *Cognition*, 158:165–176, 2017.

Brandon C. Wheeler. Production and perception of situationally variable alarm calls in wild tufted capuchin monkeys (cebus apella nigritus). *Behavioral Ecology and Sociobiology*, 64(6):989–1000, 2010.

Andrew Whiten, Jane Goodall, William C McGrew, Toshisada Nishida, Vernon Reynolds, Yukimaru Sugiyama, Caroline EG Tutin, Richard W Wrangham,

and Christophe Boesch. Cultures in chimpanzees. *Nature*, 399(6737):682–685, 1999.

Anne Wierzbicka. *English: Meaning and Culture*. Oxford University Press, USA, 2006.

Deidre Wilson and Dan Sperber. *Meaning and Relevance*. Cambridge University Press, 2012.

Heinz Wimmer and Josef Perner. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128, 1 1983.

Ludwig Wittgenstein. *Tractatus Logico-Philosophicus*. Dover Publications, 1922.

Ludwig Wittgenstein. *Philosophical investigations. Philosophische Untersuchungen*. Macmillan, Oxford, England, 1953.

Tadeusz W. Zawidzki. The function of folk psychology: Mind reading or mind shaping? *Philosophical Explorations*, 11(3):193–210, 2008.

Tadeusz W. Zawidzki. Mindshaping. In Albert Newen, Leon C. de Bruin, and Gallagher Shaun, editors, *Oxford Handbook of 4E Cognition*. Oxford University Press, forthcoming.

Tadeusz W. Zawidzki Zawidzki. *Mindshaping: A New Framework for Understanding Human Social Cognition*. A Bradford Book, 2013.