UNIVERSITÀ VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATO DI RICERCA IN FILOSOFIA

Curriculum in Discipline Filosofiche

Tesi in co-tutela con UNIVERSITAT DE BARCELONA

Corso in Cognitive Science and Language

AN EXPERIMENTAL DEFENSE OF THE CAUSAL-HISTORICAL THEORY OF REFERENCE FOR PROPER NAMES

Tutore: Dott.ssa Bianca Cepollaro

Co-tutori: Prof.ssa Genoveva Martí,

Prof. Filippo Domaneschi

Tesi di DOTTORATO DI RICERCA di Nicolò D'Agruma

matr. 017708

Ciclo di dottorato XXXVI

SSD M-FIL/05

Anno accademico UniSR 2022/2023

Anno accademico UB 2023/2024

CONSULTAZIONE TESI DI DOTTORATO DI RICERCA

Il/la sottoscritto/a / *I*	Nicolò D'Agruma
Matricola / *registration number*	017708
nato a / *born at*	Milano
il / *on*	27.03.1996

autore della tesi di Dottorato di ricerca dal titolo / *author of the PhD Thesis titled*
An experimental defense of the causal-historical theory of reference for proper names

☒ AUTORIZZA la Consultazione della tesi / AUTHORIZES *the public release of the thesis*

□ NON AUTORIZZA la Consultazione della tesi per ..... mesi /DOES NOT AUTHORIZE *the public release of the thesis for ..... months*

a partire dalla data di conseguimento del titolo e precisamente / *from the PhD thesis date, specifically*

Dal / *from* ...../...../..... Al / *to* ...../...../.....

Poiché /because:

□ l'intera ricerca o parti di essa sono potenzialmente soggette a brevettabilità / *The whole project or part of it might be subject to patentability*;

□ ci sono parti di tesi che sono già state sottoposte a un editore o sono in attesa di pubblicazione / *Parts of the thesis have been or are being submitted to a publisher or are in press*;

□ la tesi è finanziata da enti esterni che vantano dei diritti su di esse e sulla loro pubblicazione / *the thesis project is financed by external bodies that have rights over it and on its publication.*

Data / *Date* 26/01/2024 Firma / *Signature* Nicolò D'Agruma

DECLARATION

This thesis has been composed by myself and has not been used in any previous application for a degree.

All the results presented here were obtained by myself, except for:

1) Experiment 1 in Chapter 3 was conducted with Litman Huang.

2) The presentation of Experiment 3 in Chapter 4 has been re-adapted from a manuscript titled "Reference and proper names. A clash between eye-tracking and truth-value judgments", co-authored with Filippo Domaneschi, Massimiliano Vignolo and Camilo Rodríguez Ronderos, with whom Experiment 3 was conducted. The paper is under review at the time of submission of this thesis.

It is also signaled that:

1) Most of structure and content of Chapter 1 has been re-adapted and/or translated from a previously published article, whose details follow:
   D'Agruma, N. (2022). Riferimento dei nomi propri e filosofia sperimentale. *APhEx*, *25*.

2) Some arguments developed in section 2.3 have been re-adapted from a previously published article, whose details follow:
   D'Agruma, N. (2023). From the epistemic perspectives in experimental semantics to the ambiguity of proper names: Is the inference warranted? A critical notice of Jincai Li's *The referential mechanism of proper names*. *Mind & Language*, *38*(4), 1138-1146. https://doi.org/10.1111/mila.12472

All sources of information are acknowledged by means of reference.

Acknowledgments

Pursuing a PhD is both fascinating and demanding, filled with numerous challenges. Maintaining focus, self-esteem, and perseverance is tough. I could not have completed this thesis without the essential support of many, to whom I owe a great debt of gratitude.

First, I am profoundly thankful to my mentors Bianca Cepollaro, Genoveva Martí, and Filippo Domaneschi. Bianca has been a guiding light literally from day one, and I will never forget the enthusiasm with which Genoveva and Filippo accepted my proposal to co-supervise the project as soon as I contacted them. The assistance of all three, the opportunities they provided, and the doors they opened are things I will always treasure. This dissertation reflects our extensive and stimulating discussions. Their extraordinary skills as philosophers are matched only by their exceptional human qualities. I couldn't have asked for better supervisors.

My thanks also go to Università Vita-Salute San Raffaele (UniSR) in Milan and Universitat de Barcelona (UB), the institutions of my PhD. UniSR, a constant presence of my academic journey since my undergraduate days, also financially supported my PhD studies with a scholarship. Both universities, through their esteemed research centers CRESA at UniSR and LOGOS at UB, fostered dynamic spaces for intellectual engagement and personal growth. The CRESA Work-In-Progress seminars, the LOGOS Graduate Reading Group and the reading groups led by Genoveva Martí, Carl Hoefer, and Josep Macià at UB, were extremely stimulating, offering platforms for fruitful idea exchanges in supportive settings. Especially the reading group led by Genoveva, with its focus on experimental semantics, had a particularly profound impact on the content of this thesis.

I am indebted to numerous scholars in Milan, Barcelona, and globally, including leading figures in the debate of this thesis. The opportunity to meet and learn from such esteemed philosophers, whom I had only known through their writings, was a dream come true. Other scholars, while not directly connected to my specific research, have offered precious support in various ways. Thus, thanks to Andrea Bianchi, Claudia Bianchi, Francesca Boccuni, Marta Campdelacreu, Federico Cella, Daniel Cohnitz, Gregory Currie, Ryan Doran, Manuel García-Carpintero, Carl Hoefer, Markus Kneer, Paolo Labinaz, Jincai Li, Edouard Machery, Josep Macià, Carlo Martini, Nicolas Porot, Camilo Rodríguez Ronderos, Sven Rosenkranz, Martina Rosola, Elisabetta Sacchi,

Andrea Sereni, Alfredo Tomasetta, Massimiliano Vignolo, and others who contributed in numerous ways. Special acknowledgments go to Jussi Haukioja and Michael Devitt. Jussi provided essential philosophical insights with remarkable availability and patience during his visiting research period in Barcelona, greatly enriching my understanding of the topics. Michael too demonstrated exceptional willingness to help me, providing precious feedback from the first time I met him in Parma. Furthermore, the experimental-semantics gathering hosted by Michael Devitt in Hudson in June 2022, splendidly organized also by his wife Pegg Nadler, was undoubtedly the most stimulating and exciting philosophical experience I've ever had. I will never thank Michael enough for that fantastic opportunity.

The support of fellow PhD students in Milan, Barcelona, and beyond was invaluable. Sharing our experiences, encompassing both triumphs and challenges, reminded us that we're not alone in our academic efforts. Moreover, their philosophical insights significantly enhanced my work. Thus, thanks to a number of people who shared (a part of) their PhD with me: Luca Ausili, Arianna Beghetto, Federico Bina, Cristiano Calì, Duccio Calosi, Francesca Cesarano, Litman Huang, Edoardo La Ragione, Leonardo Manna, Mario Marotta, Simone Marsilio, Maurizio Mascitti, Marc Mela, Álvaro Mozota, Ilia Patronnikov, Stefano Pinzan, Paolo Polizzi, Michele Ricciotti, Niccolò Rossi, Davide Sguario, Nikolai Shurakov, Alberto Termine, Jeske Toorman, Alessandro Volpe, Eleonora Volta, and others inadvertently not mentioned.

Gratitude is extended to the attendees of my conference presentations, whose queries spurred improvements in my work and presentation skills. Moreover, hosting Edouard Machery at UniSR in March 2022 was an incredible honor: attending the lessons of one of the authors who started the debate addressed by this thesis deeply enriched my grasp of the philosophical discussions that shape this field.

I extend my gratitude to the people with whom I collaborated on experiments and whom I already mentioned: Litman Huang, Filippo Domaneschi, Massimiliano Vignolo, and Camilo Rodríguez Ronderos. Their availability, brilliance and patience made everything possible. Litman opened the door to a fascinating cross-cultural study between Italy and Hong Kong, while Filippo, Massimiliano, and Camilo introduced me to the innovative eye-tracking technique. I also thank all the participants in the experiments, particularly those who volunteered their time and helped in spreading the surveys.

This journey would not have been possible without the unwavering support of my long-time friends outside academia. Many have been with me for years, with some friendships tracing back to our earliest school days. Together, we have created a tapestry of memories, each springing vividly to mind whenever I revisit buildings and places that played an important role throughout my adolescence and life. Each of you has taught me invaluable lessons and become a part of who I am. You all know who you are, and I am profoundly grateful for your presence in my life.

Finally, my deepest gratitude goes to my family: my mum Rosaria, my dad Leonardo, my brother Federico, my sister Maria Chiara, and my uncle Giuseppe. They have always supported my philosophical pursuits, even when other paths might have seemed more pragmatic. They have been there for me in every conceivable way, shaping who I am and who I aim to become. Their embodiment of moral virtue stands as the greatest example in my life. Words simply cannot express the depth of the gratitude and love I have for them.

Abstract – English

This dissertation champions the causal-historical theory of reference for proper names by using experimental methods, engaging with the debate that Machery, Mallon, Nichols, and Stich raised in 2004 with their widely influential article "Semantics, cross-cultural style". Machery and colleagues challenged Kripke's refutation of descriptivism in *Naming and Necessity*, by presenting the "Gödel Case" to Westerners and East Asians and concluding that referential intuitions vary intra and cross-culturally.

Chapter 1 delves into the two major criticisms against Machery et al.'s work. The first concerns the ambiguities in their final prompt; the second asserts that scholars should test theories of reference against linguistic usage and not referential intuitions, as these have a meta-linguistic nature.

Chapter 2 explores the major linguistic-usage studies in the literature. While the linguistic-production data with USA participants support the causal-historical theory, the linguistic-comprehension data, based on truth-value judgments of sentences and involving both Westerners and East Asians, are less conclusive due to an epistemic ambiguity. Scholars tried to address the ambiguity with follow-up questions, which however reintroduce meta-linguistic reflections.

Chapter 3 discusses two truth-value judgment studies: Experiment 1, with Chinese and Italian participants, conducted in collaboration with Litman Huang, and Experiment 2, with Italian participants. These works aim to resolve the epistemic ambiguity by focusing on the "Jonah Case", and corroborate the causal-historical theory.

Chapter 4 presents an eye-tracking study on the Gödel Case, Experiment 3, conducted with Filippo Domaneschi, Massimiliano Vignolo and Camilo Rodríguez Ronderos. The eye-tracking methodology, bypassing explicit judgments from participants, provides direct linguistic-usage evidence and supports the Kripkean theory. Moreover, the truth-value judgment data of this experiment differ from the eye-tracking results. This divergence raises questions about the reliability of truth-value judgments and suggests the need for further research.

The dissertation concludes with Chapter 5, which summarizes the arguments and develops additional considerations on the role of empirical data in the philosophy of language.

Abstract – Italiano

Questa tesi difende la teoria storico-causale del riferimento per i nomi propri utilizzando metodi sperimentali, inserendosi nel dibattito iniziato da Machery, Mallon, Nichols e Stich nel 2004 con il loro influente articolo "Semantics, cross-cultural style". Machery e collaboratori hanno messo in discussione la confutazione del descrittivismo da parte Kripke in *Naming and Necessity*, sottoponendo il "Caso Gödel" a occidentali ed est-asiatici e concludendo che le intuizioni referenziali variano sia all'interno delle culture sia tra culture diverse.

Il Capitolo 1 esamina le due principali critiche al lavoro di Machery et al. La prima riguarda le ambiguità nella loro domanda finale; la seconda afferma che gli studiosi devono testare le teorie del riferimento basandosi sull'uso linguistico e non sulle intuizioni referenziali, poiché queste hanno una natura meta-linguistica.

Il Capitolo 2 esplora i principali studi sull'uso linguistico nella letteratura. Mentre i dati sulla produzione linguistica con partecipanti statunitensi supportano la teoria storico-causale, i dati sulla comprensione linguistica, che si fondano su giudizi sul valore di verità di enunciati e che coinvolgono sia occidentali che est-asiatici, sono meno conclusivi a causa di un'ambiguità epistemica. Gli studiosi hanno cercato di affrontare l'ambiguità con domande di follow-up, che però reintroducono riflessioni meta-linguistiche.

Il Capitolo 3 discute due studi basati su giudizi sul valore di verità: Esperimento 1, con partecipanti cinesi e italiani, condotto in collaborazione con Litman Huang, ed Esperimento 2, con partecipanti italiani. Questi lavori mirano a risolvere l'ambiguità epistemica concentrandosi sul "Caso Giona", e corroborano la teoria storico-causale.

Il Capitolo 4 presenta uno studio di eye-tracking sul Caso Gödel, Esperimento 3, condotto con Filippo Domaneschi, Massimiliano Vignolo e Camilo Rodríguez Ronderos. La metodologia di eye-tracking, evitando giudizi espliciti da parte dei partecipanti, fornisce evidenza diretta sull'uso linguistico e supporta la teoria kripkeana. Inoltre, i dati sui giudizi sul valore di verità di questo esperimento differiscono dai risultati di eye-tracking. Questa divergenza solleva domande sull'affidabilità dei giudizi di verità e suggerisce la necessità di ulteriore ricerca.

La tesi si conclude con il Capitolo 5, che riassume gli argomenti e sviluppa considerazioni aggiuntive sul ruolo dei dati empirici in filosofia del linguaggio.

Abstract – Español

Esta tesis defiende la teoría histórico-causal de la referencia para los nombres propios utilizando métodos experimentales, insiriéndose en el debate iniciado entre Machery, Mallon, Nichols y Stich en 2004 con su influyente artículo "Semantics, cross-cultural style". Machery y sus colaboradores pusieron sobre la mesa la refutación del descriptivismo elaborada por Kripke en *Naming and Necessity*, presentando el "Caso Gödel" a occidentales y asiáticos del este y concluyendo que las intuiciones referenciales varían tanto intraculturalmente como entre culturas distintas.

El Capítulo 1 examina las dos principales objeciones al trabajo de Machery et al. La primera se centra en las ambigüedades de su pregunta final; la segunda afirma que los estudiosos deben testar las teorías de la referencia basándose en el uso lingüístico, no en las intuiciones referenciales, ya que estas tienen una naturaleza meta-lingüística.

El Capítulo 2 explora los principales estudios sobre el uso lingüístico en la literatura. Mientras los datos sobre la producción lingüística con participantes estadounidenses apoyan la teoría histórico-causal, los datos sobre la comprensión lingüística, basados en juicios sobre el valor de verdad de enunciados y llevados a cabo tanto por sujetos estadounidenses como asiáticos, son menos concluyentes debido a una ambigüedad epistémica. Los estudiosos han intentado afrontar esta ambigüedad a través de preguntas de seguimiento que, sin embargo, reintroducen reflexiones meta-lingüísticas.

El Capítulo 3 discute dos estudios basados en juicios de verdad: Experimento 1 con participantes chinos e italianos, realizado en colaboración con Litman Huang, y Experimento 2 con participantes italianos. Estos trabajos intentan resolver la ambigüedad epistémica centrándose en el "Caso Jonás", y corroboran la teoría histórico-causal.

El Capítulo 4 presenta un estudio de registro visual sobre el Caso Gödel, Experimento 3, realizado con Filippo Domaneschi, Massimiliano Vignolo y Camilo Rodríguez Ronderos. La metodología de registro visual, evitando juicios explícitos por parte de los participantes, proporciona evidencia directa sobre el uso lingüístico y apoya la teoría kripkeana. Además, los datos de juicios sobre valor de verdad de este experimento difieren de los resultados de registro visual. Esta divergencia plantea preguntas sobre la fiabilidad de los juicios de verdad y sugiere la necesidad de una investigación posterior.

La tesis concluye con el Capítulo 5, que resume los argumentos y presenta consideraciones adicionales sobre el rol de los datos empíricos en la filosofía del lenguaje.

Table of contents

Abbreviations

DT	Descriptivist Theory
CHT	Causal-Historical Theory
MMNS	Machery, Mallon, Nichols and Stich
MOD	Machery, Olivola and de Blanc
MSD	Machery, Sytsma and Deutsch
RI	Referential Intuition
TVJ	Truth-Value Judgment
EP	Elicited Production

Introduction

Language is crucial in our life. People use language, among other things, to talk about entities of the world and attribute properties to them. This is possible by virtue of the referential capacity of certain linguistic items. Some words stand for entities, typically extra-linguistic ones. For example, when a speaker utters "Bachelors are happy", what she does is to refer to bachelors by the word "bachelors" and attribute them a property, that of being happy. One of the main questions in the philosophy of language pertains to how it is possible that words stand for things, namely, that words refer to entities. A natural answer is that a word refers to what it does by virtue of its definition, namely a description through which that specific word refers to an entity or a group of entities. For example, the definition of "bachelor" is "unmarried man", and the word "bachelor" picks out bachelors by virtue of that definition. That is, the group of bachelors is what the term "bachelor" stands for because those people satisfy the description "unmarried man" that speakers associate with the term. After all, when a person does not know what a word stands for, she looks for the definition of that word in a dictionary, which will provide the description that determines the domain of application of that term.

However, the above picture of how words refer to things is necessarily incomplete. If that constituted the entire explanation of referential relations, it would ultimately be impossible for any word to refer to things in the world. What composes the definition of a word, as documented in any dictionary, is a description, which in turn consists of words. These words will serve their role of establishing a referential relation if and only if they entertain some referential relations with things and portions of the world. For example, the description "unmarried man" establishes a referential link for "bachelor" only if "unmarried" and "man" successfully refer to entities or portions of the world. Therefore, these words too need a descriptive definition that mediates their relation with the world. However, these definitions amount to words, which in turn need a definition to fulfill their referential role. It thus seems that language becomes ensnared in a circular pattern and thus is cut loose from reality: words call for other words, which in turn call for more words, and at some point the chain of reference inevitably leads back to its starting term.

Language would never surpass its own domain and get anchored to the extra-linguistic world. Russell (1918/1919) stresses the point:

> A dictionary professes to define all words in the language by means of words in the language, and therefore it is clear that a dictionary must be guilty of a vicious circle somewhere.
> (Russell, 1918/1919, p. 22, reference to the republished text)

Devitt and Sterelny (1999) too emphasize a similar aspect:

> There must be some terms whose referential properties are not dependent on others. Otherwise, language as a whole is cut loose from the world. Description theories, which explain one part of language in terms of another, can give no clue as to how, ultimately, language is referentially linked to reality. These theories pass the referential buck. But the buck must stop somewhere.
> (Devitt and Sterelny, 1999, p. 60)

One of the most groundbreaking works in the philosophy of language is *Naming and Necessity*, a series of lectures that Kripke held at Princeton University in 1970, later compiled into a book (1972/1980).[1] Among the several crucial contributions that Kripke's work offers is the individuation of two categories of terms that would refer without the mediating role of descriptions. Those two categories are proper names, that is expressions like "Barack Obama", "Uranus" and "Italy", and natural-kind terms, that is expressions like "gold", "tiger" and "water". Those words would constitute at least a part of the explanation of how language is not cut loose from the world. The present dissertation focuses on proper names. Kripke contends that what enables a proper name to refer to a specific entity of the world is not the description that speakers associate with the term, but rather a chain of uses that connects that name with a specific entity upon whom someone originally bestowed the name itself. Kripke's view is called "causal-historical theory" and contrasts with the descriptivist theory, defended by Frege (1982),

[1] Henceforth, I shall reference this work by using the publication year of the second edition (Kripke, 1980), as scholars commonly cite it.

Russell (1905, 1918/1919, 1919) and Searle (1958), according to which names refer through descriptions.

In order to prove his point, Kripke conceives – among the other ones – a very famous thought experiment, called the "Gödel Case", and wonders about the referential intuition that it elicits – a referential intuition is a judgment as to whom a speaker talks about when using a term. The Gödel Case describes a counterfactual situation in which Gödel is not the discoverer of the incompleteness theorem and stole the results from another mathematician, Schmidt (Kripke, 1980, pp. 93-4). In that scenario, anyone associates the name "Gödel" with the only description "The discoverer of the incompleteness theorem". Kripke claims that the common referential intuition on the Gödel Case is that a speaker using "Gödel" in such a circumstance talks about Gödel, not Schmidt, even though Schmidt satisfies the description that the speaker associates with the name. That thought experiment would show that what determines the name's ability to refer is not the description that speakers associate with it, but some other non-descriptive mechanism. According to Kripke, that mechanism consists of a chain of communication: roughly speaking, a person introduces the name to talk about a specific entity. The name spreads within a linguistic community and whoever uses the name will refer to the entity for which it was initially introduced, irrespective of whether that entity satisfies the description that the speaker may possibly associate with that name. Kripke sketches the mechanism in one of the most quoted passages of *Naming and Necessity*:

> Someone, let's say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman is a famous physicist. A certain passage of communication reaching ultimately to the man himself [Feynman] does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely [or even though the speaker's description identifies someone uniquely, but that person is not Feynman]. […] a chain of communication going back to Feynman himself has been

> established, by virtue of his [the speaker's] membership in a community which passed the name on from link to link […].
> (Kripke, 1980, p. 91)[2]

As I stressed, Kripke's thought experiment on the Gödel Case relies upon a supposedly shared intuition, namely that a speaker, when using the name "Gödel" in that scenario, refers to Gödel and not to Schmidt. As a matter of fact, that intuition proved to be shared within philosophical circles. [3] However, one may wonder what philosophical consequences would follow if that intuition failed to find widespread agreement among the more general population. Without a reasoned principle to prefer philosophers' intuitions over laypeople's, it is narcissistic to dismiss the disagreement possibly arising from the latter group. One of the critical aims of experimental philosophy is to empirically investigate whether the intuitions to which philosophers appeal in order to develop their arguments are really universal and, if that is not the case, what the implications are. Typically, experimental philosophers present laypeople with vignettes modeled upon the thought experiments that philosophers conceive. The different options at the end of the vignette capture the possible intuitive reactions that participants may have as regards those cases.

In 2004 Machery, Mallon, Nichols and Stich (MMNS) publish a groundbreaking experimental article, "Semantics Cross-Cultural Style". They provide data that would corroborate two conclusions. First, the referential intuitions on the Gödel Case vary across Westerners and East Asians: MMNS's American sample is more inclined to express an intuition in line with the Kripkean theory of names, while the Chinese one is more inclined toward the descriptivist view. Moreover, even within each of the two

[2] The quoted passage focuses on the so-called Feynman Case, which highlights that a speaker, when using a name, refers to the entity at the end of its communication chain even if the speaker lacks a uniquely identifying description (e.g., if she associates the name "Feynman" with the general description "some physicist"). However, this introduction focuses on the scenario in which a speaker associates the name with a description uniquely picking out an individual different from the one at the end of the communication chain, as that is the scenario that the Gödel Case describes and that virtually all the experimental literature targets.

[3] Beebe and Undercoffer (2016, p. 342) and Häggqvist and Wikforss (2015, p. 131) label the Kripkean view of names (and natural kind-terms) as the "orthodoxy" in the philosophy of language. Devitt (2020, p. 438) too writes: «given the response to *Naming and Necessity* in the literature, we can reasonably assume that philosophers, particularly philosophers of language, generally share Kripke's referential intuitions. And this assumption has received some experimental support: in an experiment on a Gödel case, the intuitions of semanticists and philosophers of language were decisively antidescriptivist (Machery 2012a)» (*ibidem*). I discuss Machery (2012a) in detail in section 1.4.1. Also Sytsma and Livengood (2011) present one of their vignettes to philosophers, collecting a high proportion of causal-historical intuitions, though not as high as the one reported in Machery (2012a). I explore Sytsma and Livengood's results in section 1.3.2.

cultural backgrounds, a non-negligible variation between the two theories emerges: the answers of the American sample are not consistently Kripkean and those of the Chinese sample are not consistently descriptivist. These results constitute a challenge to Kripke's supposed refutation of the descriptivist view of proper names, as the intuition upon which he relies for the Gödel Case does not seem to be universal. MMNS's paper evoked a number of reactions and generated an impressive body of literature. This dissertation endorses the experimental approach to study the reference of proper names and thus locates itself in the experimental debate that MMNS (2004) elicited. I ultimately defend the Kripkean view of names.

In Chapter 1, I present the two main theories on the reference for proper names, the descriptivist view and the causal-historical view, and discuss MMNS's experimental challenge against Kripke's refutation of the descriptivist theory. I present the criticisms that scholars raise against MMNS's work, by dividing them into two categories: the ones strictly related to the experimental design and the ones of a more theoretical nature. The former criticisms pertain to some ambiguities that would affect the referential intuitions that MMNS collect (Ludwig, 2007; Deutsch, 2009; Sytsma and Livengood, 2011). The latter criticisms pertain to the kind of data upon which Kripke's arguments would rely (Deutsch, 2009, 2010, 2015; Martí, 2009, 2012, 2014, 2020) and the kind of empirical research – if any – that philosophers need to conduct to test theories of reference. Martí (2009, 2012, 2014, 2020) and Devitt (2011, 2012b, 2015b) argue that scholars should test theories of reference against linguistic usage and not referential intuitions. Moreover, in Chapter 1 I delve into the relatively less extensive experimental literature on the Jonah Case, which is another scenario that Kripke uses to support his own theory of names. All the information that people inherited about the man called "Jonah" is the Biblical one, which is arguably true of no real individual: the Bible describes the feats of a man who, for example, would have lived in the stomach of a big fish for three days. Therefore, Kripke highlights that, according to the descriptivist view, the name "Jonah" is devoid of reference. However, that prediction is inconsistent with a seemingly plausible situation: Jonah may have really existed (i.e., "Jonah" refers to a real man), but we only have false information about that person. The causal-historical theory can accommodate such a situation, as the view implies that the name refers to the individual at the end of the communication chain, irrespective of whether the information that speakers associate

with the name is true or false of that person. Although MMNS test the Jonah Case as well, that scenario has received comparatively less attention than the Gödel one, primarily because MMNS fail to find the expected cross-cultural variation between Americans and Chinese.

In Chapter 2, I argue that Martí and Devitt are right that scholars should test theories of reference against linguistic usage and not referential intuitions. Moreover, I endorse Cohnitz and Haukioja's (2013, 2015) point that linguistic usage encompasses not only production but also comprehension. Therefore, I develop an in-depth analysis of what I take to be the three main linguistic-usage studies: Devitt and Porot (2018), Domaneschi and Vignolo (2020) and Li (2021). As regards Devitt and Porot (2018), I express some reservations about their critical examination of MMNS's (2004) vignettes and their theoretical analysis of what they term the "New Meaning Objection". That said, the way they experimentally tackle that objection is successful and their linguistic-production data with American participants successfully support the Kripkean theory of names. However, it remains crucial to examine the other side of linguistic usage as well, namely comprehension: if also the data stemming from that side of use align with the same direction, the causal-historical theory would gain even more solid corroboration. This is what Domaneschi and Vignolo (2020) and Li (2021) do with their truth-value judgment tasks, testing Italian participants (Domaneschi and Vignolo) and American and Chinese ones (Li). I argue that their works, compared to other studies collecting truth-value judgments (e.g., Machery, Olivola and de Blanc, 2009), have the merit of identifying a possible epistemic ambiguity that parallels the one that Sytsma and Livengood (2011) identify in the context of referential intuitions. However, the way Domaneschi and Vignolo on the one hand and Li on the other try to control for the epistemic ambiguity is problematic, primarily because they use follow-up questions, which reintroduce the meta-linguistic reflections Martí (2009) warns researchers against. I also argue that, although neither Domaneschi and Vignolo nor Li emphasize this aspect, the critical sentence at the center of their truth-value judgment tests has the merit of controlling for the potential attributive reading of names, a confounding factor that Heck (2018) and Vignolo and Domaneschi (2022) highlight.

In Chapter 3, I present Experiment 1 and Experiment 2. Inspired by a methodology that Islam and Baggio (2020) adopt, Experiment 1 and Experiment 2 try to control for the

epistemic ambiguity by placing emphasis on the Jonah Case. Experiment 1, coauthored with Litman Huang, presents some explorative TVJ data on the Jonah Case with an Italian sample and a Chinese one. The data's preliminary nature is due to certain potential design issues in the study. Nonetheless, Experiment 1 corroborates CHT as the correct theory for both samples: only a limited number of the Italian and Chinese participants opt for the TVJ "False", which is a significant finding given that in the Jonah Case that TVJ can only express a descriptivist understanding of the name. Experiment 2, which involves only Italian participants, tests linguistic comprehension by directly collecting data on both the Gödel Case and the Jonah Case. In this way, it is possible to control for the epistemic ambiguity while circumventing the reliance on follow-up questions. I take the results to corroborate various theses. First, the causal-historical view is the correct theory of reference for proper names. Second, the attributive interpretation of names is a confounding factor that indeed influenced the previous Gödel Case literature on truth-value judgments and referential intuitions. Third, the lingering purportedly descriptivist evidence that the Gödel Case seems to elicit, even after one controls for the attributive reading of names, corresponds to the extent to which participants adopt an epistemic perspective that is unsuitable to compare theories on reference.

In Chapter 4, I present Experiment 3, a linguistic-comprehension eye-tracking experiment on the Gödel Case, which Filippo Domaneschi, Massimiliano Vignolo and Camilo Rodríguez Ronderos and I conducted with a sample of Italian participants. The theoretical background for the study is this. As an off-line measure, a truth-value judgment test comprises two stages: first, a comprehension phase in which the participant understands the proposition that a specific sentence expresses; second, the formulation of a truth-value judgment that is supposed to shed light upon that proposition. In between the two stages, some meta-linguistic theoretical reflections – the same ones that contaminate referential intuitions – may intervene, by finally invalidating the resulting truth-value judgment. Therefore, truth-value judgments would face limitations stemming not only from the epistemic ambiguity of the truth predicate, but also from the possibility that participants meta-linguistically reflect on theories of reference as applied to proper names. I argue that, while on the one hand the epistemic ambiguity is a real threat, on the other hand explicit meta-linguistic reflections on theories of reference are arguably more unlikely to emerge in the context of a truth-value judgment test as opposed to a referential

intuition test, as only the latter explicitly asks the participant to consider the name "Gödel" (or any name around which the vignette revolves). Nevertheless, it is unquestionably beneficial to use also an online, direct methodology that complements the truth-value judgment results and provides the benchmark data as regards linguistic comprehension. Building on these considerations, as said, we conducted our eye-tracking experiment on the Gödel Case. In doing that, we followed Cohnitz and Haukioja's (2015) and Cohnitz's (2015) idea to resort to the eye-tracking technique to test the reference of proper names. The eye-tracking methodology is an online one, as it investigates the comprehension process while it unfolds, namely, while the participant hears the critical sentence. The results corroborate the causal-historical view as the correct theory of reference. Furthermore, during the eye-tracking experiment, we collected truth-value judgment data on the same Gödel scenarios. The truth-value judgment results strongly align with the descriptivist prediction and thus diverge from the causal-historical eye-tracking findings substantially: such a discrepancy seems to vindicate Martí's concern as to the reliability of truth-value judgments. However, the source of the unreliability is not clear. The truth-value judgments that we gather diverge from those of Experiment 2 and more in general from the previous literature with Western participants. Therefore, the seemingly descriptivist truth-value judgments of our sample may stem from the inclusion of the truth-value judgment task within an eye-tracking experiment. Some features related to the eye-tracking design may have impacted the truth-value judgment results, maybe combining with the epistemic ambiguity or the meta-linguistic reflections that Martí stresses, albeit in ways that still lack precise identification. Therefore, generally speaking, truth-value judgments might be overall reliable, although some specific features of our eye-tracking design may have made them exceptionally unreliable in the context of our study.

In Chapter 5, I draw the conclusions of my work. The linguistic-comprehension data that I present in my dissertation corroborate the causal-historical theory for Western speakers. Moreover, the data of Experiment 1 tentatively corroborate that theory also for a Chinese sample or at least suggest that it would be naïve to conclude that people belonging to that demographic are descriptivist users of names. The seemingly descriptive evidence that the previous truth-value judgment literature elicits as regards the Gödel Case arguably stems from various confounding factors: the possible attributive

understanding of the name, an inherent epistemic ambiguity and, perhaps, explicit theoretical reflections on names. I therefore reject those views, like the one that Li (2023a) defends on an experimental basis, according to which proper names are ambiguous between two referential mechanisms – a causal-historical one and a descriptivist one. Kripke's refutation of descriptivism successfully withstands the experimental challenge. The circumstance that the experimental data leads to the same conclusion that philosophers have traditionally welcomed, namely that Kripke successfully refuted descriptivism, might seem to somehow downsize the philosophical value of the experimental research that has stemmed from MMNS (2004). I strongly disagree: MMNS (2004) has had the invaluable merit of emphasizing the necessity for scholars to test theories of reference against experimental data. Absent this empirical verification, those theories would rely upon purely conjectural grounds. Such an infirm basis would be problematic not only for theories of reference *per se*, but also for those views, for example in the field of epistemology or metaphysics, that depend on the soundness of some specific theory of reference (Mallon et al., 2009).

1 Proper names and experimental philosophy

1.1 Theories on the reference of names

Proper names are expressions like "John", "Margaret", "Paris" and "The Starry Night". Theories on the reference of proper names explain what determines that an object *o* is the bearer of a proper name N as used by a speaker S. Traditionally, two theories have been at the center of the debate on the reference of proper names: the descriptivist theory (DT) and the causal-historical theory (CHT).

According to DT, proposed by Frege (1892)[4], Russell (1905, 1918/1919, 1919) and Searle (1958), the referent of a name N is the object *o* that uniquely satisfies the description associated with N or satisfies most of the weighted content in the cluster of descriptions associated with N. Absent an object that uniquely satisfies the description or that is the best satisfier of the cluster, N is devoid of referent. N lacks a referent also in the case multiple entities satisfy the description or most of the cluster of descriptions.

Kripke develops CHT in *Naming and Necessity* (1980)[5]. Someone bestows a name N upon an object *o* by an act of baptism. Progressively, the name spreads within a linguistic community thanks to a chain of uses, in which each speaker learns N from another. In this way, N reaches a subject S. By using N, S will refer to the entity for which N was initially introduced, i.e. *o*, irrespective of whether *o* satisfies the description (or the cluster of descriptions) that S associates with N. Therefore, when using N, S refers to *o* also in cases where the description is false of *o* or not precise enough to uniquely pick out *o*. Moreover, even in those situations where the description is true of *o*, it is not that circumstance that makes *o* the referent of N. Rather, it is the fact that *o* is at the far end of the chain of uses of N.[6]

The Kripkean theory is called "causal-historical". However, as Martí stresses (2012, 2015), it is not clear where the supposedly "causal" element in the chain of communication lies. Rather, the hallmark of the theory seems to be its social character.

[4] Dummett (1973) objects that a thorough examination of Frege's work reveals that he does not fully adhere to descriptivism concerning proper names. However, as Kripke (1979) notes, «the philosophical community has generally interpreted Fregean senses in relation to descriptions» (p. 271, n. 3). I will follow the traditional interpretation of Frege's view. In any case, the precise philological details of Frege's stance do not have any substantial bearing on the arguments in this dissertation.

[5] In one of his earlier publications, Kripke (1971) already presents some fundamental aspects of CHT. Donnellan (1970) too puts forth ideas that bear some resemblance to Kripke's.

[6] In the dissertation, I will use "referent" to talk about the object to which the name refers and "reference" to talk about the name's capacity to refer, which includes the mechanism through which it refers – descriptivist or causal-historical.

The referent of a name depends on how other members of the linguistic community previously introduced and used the name. Kripke himself, despite introducing the expression "causal chain" (Kripke, 1980, p. 93), sometimes uses this adjective only in scare quotes (p. 59, n. 22) or simply speaks of "chain of communication" (p. 91). That said, in this work I will adhere to the current terminology and use the expression "causal-historical".

In order to corroborate his proposal, Kripke presents some thought experiments, namely some «nonactual situations[s]» that carry philosophical implications relevant to the topic under discussion, in this case proper names (Machery, 2011a, p. 194). Kripke maintains that CHT, in contrast to DT, predicts the *referential intuitions* (RIs) that speakers exhibit about those cases. RIs are quick and spontaneous judgments regarding what a speaker talks about when using a term, like a proper name. Two central cases that Kripke presents are the *Gödel Case* and the *Jonah Case*. The Gödel Case is the following:

> Suppose that Gödel was not in fact the author of this theorem [the theorem of incompleteness]. A man named "Schmidt", whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the view in question [the descriptivist view], then, when our ordinary man uses the name "Gödel", he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description, "the man who discovered the incompleteness of arithmetic". […] since the man who discovered the incompleteness of arithmetic is in fact Schmidt, [then according to DT] we, when we talk about "Gödel", are in fact always referring to Schmidt. But it seems to me that we are not. We simply are not.
> (Kripke, 1980, pp. 83-4)

According to Kripke, the common intuition is that, in such a scenario, a speaker S using "Gödel" refers to Gödel and not Schmidt, even though S associates "Gödel" only with "The discoverer of the incompleteness theorem". CHT accounts for this intuition because Gödel, not Schmidt, is at the end of the causal-historical chain.

In the Jonah Case, Kripke stresses that the only information that speakers associate with the name "Jonah" is the one that the Bible provides. According to that narrative,

Jonah is a prophet who, astonishingly, resided within the belly of a large sea creature for a duration of three days, alongside accomplishing other remarkable feats. Clearly, no individual satisfies that description to a reasonable extent. Hence, according to DT, the name "Jonah" is devoid of reference. However, Kripke explains that such an implication is problematic:

> Suppose that someone says that no prophet ever was swallowed by a big fish or a whale. Does it follow, on that basis, that Jonah did not exist? There still seems to be the question whether the Biblical account is a legendary account of no person or a legendary account built on a real person. In the latter case, it's only natural to say that, though Jonah did exist, no one did the things commonly related to him.
> (Kripke, 1980, p. 67)

CHT accounts for the intuition that Kripke presents. If Jonah is a historically existing figure for whom our available information is entirely false (except for trivial predicates such as "being a man"), CHT entails that "Jonah" has a referent, namely the individual who people in the past originally called "Jonah".[7] Our uses of the names are connected to that individual through a chain of communication that has spread over time.[8]

Kripke supports CHT by relying upon what scholars currently call the "method of cases" (Weinberg, 2017; Horvath and Koch, 2020). He presents some hypothetical or actual scenarios, which constitute counterexamples against a theory (in this case, DT) and instead are consistent with some other theory (in this case, CHT). These cases hold

[7] Throughout this dissertation, I will use the expression "originally called" to talk about the individual whom a vignette or scenario presents as causally-historically connected to the critical name. This terminology is questionable, as the descriptivist referent is nothing but the name's "original bearer" according to DT. In other words, both DT and CHT agree that there is an "original bearer of a name", but they differ in their understanding of the mechanism by which an individual becomes the name's bearer. The use of the expression "original bearer" to talk about the individual causally-historically connected to the name stems from the way the Gödel and Jonah scenarios are usually framed. They first present a certain individual X having the name N and then explain that people started associating N with a description that is not satisfied by X, but by Y or by no entity at all. Therefore, there is a chronological sense in which X is the first bearer of N, and thereby N's "original bearer". This rationale underlies my decision to label the causal-historical referent as the name's "original bearer". That said, nothing essential hinges on whether one accepts my terminological convention.

[8] One may wonder whether the Jonah Case, strictly speaking, is a thought experiment. While the Gödel scenario asks the reader to imagine a counterfactual and thus nonactual situation (one in which Gödel did not discover the theorem), the Jonah Case does not revolve around a nonactual scenario. Rather, the case describes what is plausibly the truth, namely that there existed no individual who resided within the belly of a large sea creature for three days. However, the case also asks the reader to assume that an individual originally called "Jonah" existed (i.e., that the Bible account is built upon a real individual), *irrespective of whether* that is *actually* the case (there is a debate on the actual existence of such a figure, as Kripke himself (2013, p. 28, n. 27) notes). To that extent, one may consider the Jonah Case too as a thought experiment. Anyhow, establishing whether the Jonah Case is a thought experiment is not essential: what matters is how Kripke uses it to elicit intuitions that are purportedly inconsistent with DT.

significance in testing philosophical theories by eliciting intuitions that carry evidential weight. If a theory predicts that *p* and one has the intuition that *p*, then that is evidence supporting the theory. Instead, if one has the intuition that *not-p*, that is evidence against the theory. While presenting the Kripkean arguments, I relied upon this interpretation of the method of cases. For example, DT predicts that, in the above-described scenario, "Gödel" refers to the real discoverer of the theorem. Therefore, the purported intuition that "Gödel" refers to the thief serves as evidential support against DT and in favor of CHT. Some scholars criticized this interpretation of the method of cases, according to which their evidential value consists in their ability to elicit intuitions. I will discuss this criticism later (sections 1.4.2 and 1.4.3.1), when I present Deutsch's and Martí's positions.

1.2 The experimental challenge

Experimental philosophy is a branch of philosophy that applies the experimental methods typical of social sciences to philosophical investigation. Weinberg (2017) provides a perspicuous description of the theoretical rationale guiding experimental philosophers:

> I take x-phi [i.e., experimental philosophy] to be constituted most broadly by a thesis, and a research motivation as a clear consequence of it. The thesis: philosophical practice has substantially deeper empirical commitments than generally presupposed, in our explicitly stated theories but also in our methods, our practices with our evidential sources, and our modes of inference. The research motivation, accordingly, is for us philosophers to take those commitments seriously, and to evaluate where they are – or are not – adequately fulfilled. Perhaps we do so working on our own, or perhaps in interdisciplinary collaborations, but either way, we experimental philosophers are motivated by a desire to play a direct role in taking on this methodological responsibility.
> (Weinberg, 2017, p. 161)

One of the movement's goals is to judge whether the intuitions of the philosophical community are a reliable source of evidence (Cohnitz and Häggqvist, 2009; Genone, 2012; Alexander et al., 2010; Nado; 2016; Knobe and Nichols, 2017; Weinberg, 2017; Horvath and Koch, 2020). This critical approach also engaged the debate on the reference

of proper names. The experimental research contested the assumption that the Kripkean intuitions regarding the Gödel and Jonah Case are universally shared. While the Kripkean intuitions have enjoyed considerable popularity within the philosophical community, this widespread acceptance does not provide sufficient grounds to assume that they are universally representative. Philosophers' intuitions may simply mirror a strict community of speakers. It is an empirical investigation that must rather substantiate the supposed assumption of universality.[9] More radically, philosophers' reported intuitions may even fail to appropriately reflect their genuine reaction to the famous thought experiments on proper names that circulate within the philosophical community. Rather, the agreement among philosophers could just be the result of «academic sociology» (Tobia et al., 2020, p. 23). A certain philosopher introduces a specific intuition into the debate and a group of like-minded philosophers happen to share that intuition. That initial agreement is sufficient to facilitate a range of insular intellectual pursuits, that is, «to allow for a range of thriving intramural sports among believers» (Cummins, 1998, p. 116). However, unfortunately, «[t]hose who do not share the intuitions are simply not invited to the games» (*ibidem*). The following considerations that Lewis (1994) develops as regards the philosophers' intuitions on Putnam's (1975) Twin Earth Case are conceptually extendable to the thought experiments on proper names:

> When we hear that XYZ off on Twin Earth fits many of the conditions in the cluster we are in a state of indecision about whether it deserves the name "water". […] When in a state of semantic indecision, we are often glad to go either way, and accommodate our usage temporarily to the whims of our conversational partners […]. So if some philosopher, call him Schmutnam, invites us to join him in saying that the water on Twin Earth differs in chemical composition from the water here, we will happily follow his lead. And if another philosopher, Putnam (1975), invites us to say that the stuff on Twin Earth is not water – and hence that Twoscar does not

[9] Experimental philosophy has a so-called negative program and a positive one (Weinberg, 2017). Experimental work falls under the "negative" label when its aim is to highlight the deficiencies of the method of cases and thus downsize the evidential value of the intuitions that those cases elicit. For example, if apparently non-philosophically relevant factors such as the order of presentation of the options or the vignette's fonts happen to influence people's reactions to the cases, then one may question the utility of those cases to develop a theory. Experimental work falls under the positive side of the project when the experimentalist collects data with the explicitly constructive goal of testing and thus construing a philosophical theory. Arguably, these two programs are intricately intertwined, representing different phases of the same philosophical enterprise (*ivi*, p. 163). During the negative phase, the experimentalist shows where the limits of the method of cases lie; during the positive phase, the experimentalist controls for those problems in order to ensure that the method of cases yields the right data for philosophical theorizing.

> believe that water falls from the clouds – we just as happily follow his lead. We should have followed Putnam's lead only for the duration of that conversation, then lapsed back into our accommodating state of indecision. But, sad to say, we thought that instead of playing along with a whim, we were settling a question once and for all. And so we came away lastingly misled.
> (Lewis 1994, p. 424)

As a consequence, as Turri writes (albeit in the context of a discussion on the so-called "Gettier intuition"), «the costs of challenging it [that intuition] are high. You may assume without penalty that [this specific] intuition is correct, but you must argue mightily to defend the contrary assumption» (Turri, 2016, p. 340). Therefore, as said, the supposed agreement among philosophers would just be the result of academic sociology and conformism, which may have obscured the genuine intuitions of philosophers and led them to «mischaracterize[…] their own intuitive reaction to some key cases» (*ivi*, p. 345). That is, philosophers may «lack unproblematic transparent access to their intuitions» (Turri, 2018, p. 216).[10]

The necessity to justify empirically the supposed universality of the Kripkean intuition defended by philosophers and the possibility that the philosophers' agreement may even just be the outcome of academic sociology define the conceptual framework that grounds Machery, Mallon, Nichols and Stich's (hereinafter, MMNS) 2004 groundbreaking article, "Semantics, cross-cultural style". This paper marks the inception of the empirical investigation into the reference of proper names. As the title suggests, MMNS hypothesize that RIs vary depending on demographic factors, specifically across Westerners and East Asians. According to MMNS, a body of research in cognitive psychology, such as Norenzayan et al. (2002), Nisbett (2003) and Nisbett et al. (2011), suggests that Western individuals, when engaging in explanatory processes, demonstrate a higher inclination toward employing causal frameworks compared to their East Asian

[10] It is possible to object that the sociological explanation of the agreement among philosophers is not applicable to the specific case of proper names. DT was the main "theory on the market" before Kripke's lessons: «according to the received history, until Saul Kripke's *Naming and Necessity* (1972/1980), pretty much everyone was a descriptivist about reference for proper names and natural kind terms» (Dacey and Mallon, 2016, p. 372). Therefore, the supposed academic conformism should have led the audience to resist Kripke's novel intuitions in favor of upholding the well-established descriptivist ones. The proponent of the sociological explanation may counter that sociological dynamics within the academia extend beyond mere conformism and include also other factors, such as the desire to renew an otherwise outdated debate. These additional elements may explain a shift in the intuitions prevailing within the philosophical community. I will refrain from discussing the details of the possible moves and countermoves here.

counterparts. In light of this, MMNS predict that Westerners will display intuitions in line with the Kripkean theory, given its purported causal character. Conversely, East Asians are expected to exhibit intuitions that are less aligned with CHT.

As argued in section 1.1, the Kripkean theory does not clearly encompass a causal element. Hence, following Martí (2012), one may object that MMNS's prediction lacks a solid foundation. Moreover, even when granting the inclusion of a causal element in the Kripkean framework or considering a more causally-marked variant of the theory, as the ones that Devitt (1981) and Devitt and Sterelny (1999) elaborate, Martí insists that MMNS's prediction remains not firmly grounded. Martí objects that MMNS misinterpret Nisbett et al. (2001), Nisbett (2002) and Norenzayan et al. (2002): MMNS would ascribe them a view that they do not support. Martí also considers Choi et al. (1999), a research specifically examining variations in how East Asians and Westerners make judgments about causation, and argues that the authors of these studies do not assert that Westerners engage in reasoning that is more causally oriented than East Asians. Rather, the causal reasoning of the former tends to center on individual objects, while that of the latter places greater emphasis on contextual factors. Therefore, it is not clear why East Asians should be less inclined to adopt causal reasoning *tout court*, as MMNS suggest. Rather, as also Ostertag (2013) and Devitt (2012b) highlight, Chinese should express intuitions in line with CHT, given the social character of Kripke's theory and the Chinese's tendency to prioritize contextual and relational aspects.

MMNS test 31 American undergraduates from Rutgers University (USA) and 40 from Hong Kong University (China). Participants are laypeople, namely speakers who lack any specialized philosophical expertise. MMNS present two scenarios modeled after the Gödel Case to each experimental group. One case, which I will call *Gödel 1*, uses a Western name ("Gödel"). The other, which I will call *Gödel 2*, uses an Asian name ("Tsu Ch'ung Chih"). The authors also present two other scenarios, modeled after the Jonah Case. Again, one case, which I will call *Jonah 1,* uses a Western name ("Attila"). The other case, which I will call *Jonah 2*, uses an Asian name ("Chan Wai Man"). MMNS present the scenarios in English, given that the Hong Kong University employs English as its primary language of teaching and thus students arguably possess a proficient

command of the language.[11] Each scenario ends with a query to which the two classical theories of reference provide distinct responses.

Presented below is Gödel 1:

> Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called "Schmidt", whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus, he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name "Gödel" are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel.
>
> When John uses the name "Gödel", is he talking about:
> (a) the person who really discovered the incompleteness of arithmetic? or
> (b) the person who got hold of the manuscript and claimed credit for the work?
> (MMNS, 2004, p. B6)

DT predicts (a), because the description that John associates with "Gödel" is "The discoverer of the theorem of incompleteness". CHT predicts (b), because the individual at the end of the name's causal-historical chain is the one who got hold of the manuscript.

This is Gödel 2:

> Ivy is a high-school student in Hong Kong. In her astronomy class she was taught that Tsu Ch'ung Chih was the man who first determined the precise time of the summer and winter solstices. But, like all her classmates, this is the only thing she has heard about Tsu Ch'ung Chih. Now suppose that Tsu Ch'ung Chih did not really make this discovery. He stole it from an astronomer who died soon after making the discovery. But the theft remained entirely undetected and Tsu Ch'ung Chih became

[11] Unless otherwise specified, the experimental materials discussed in this dissertation are in English.

famous for the discovery of the precise times of the solstices. Many people are like Ivy; the claim that Tsu Ch'ung Chih determined the solstice times is the only thing they have heard about him.

When Ivy uses the name "Tsu Ch'ung Chih", is she talking about:
(a) the person who really determined the solstice times? Or
(b) the person who stole the discovery of the solstice times?
(MMNS, 2004, pp. B9-B10)

The predictions that the two theories make are the same as those in Gödel 1.

As regards the Jonah cases, for the sake of brevity I will report only Jonah 1 (Jonah 2 follows an analogous structure).

In high-school, German students learn that Attila founded Germany in the second century A.D. They are taught that Attila was the king of a nomadic tribe that migrated from the east to settle in what would become Germany. Germans also believe that Attila was a merciless warrior and leader who expelled the Romans from Germany, and that after his victory against the Romans, Attila organized a large and prosperous kingdom.
Now suppose that none of this is true. No merciless warrior expelled the Romans from Germany, and Germany was not founded by a single individual. Actually, the facts are the following. In the fourth century A.D., a nobleman of low rank, called "Raditra", ruled a small and peaceful area in what today is Poland, several hundred miles from Germany. Raditra was a wise and gentle man who managed to preserve the peace in the small land he was ruling. For this reason, he quickly became the main character of many stories and legends. These stories were passed on from one generation of peasants to the next. But often when the story was passed on the peasants would embellish it, adding imaginary details and dropping some true facts to make the story more exciting. From a peaceful nobleman of low rank, Raditra was gradually transformed into a warrior fighting for his land. When the legend reached Germany, it told of a merciless warrior who was victorious against the Romans. By the eighth century A.D., the story told of an Eastern king who expelled the Romans and founded Germany. By that time, not a single true fact remained in the story.
Meanwhile, as the story was told and retold, the name "Raditra" was slowly altered: it was successively replaced by "Aditra", then by "Arritrak" in the sixth century, by

> "Arrita" and "Arrila" in the seventh and finally by "Attila". The story about the glorious life of Attila was written down in the eighth century by a scrupulous Catholic monk, from whom all our beliefs are derived. Of course, Germans know nothing about these real events. They believe a story about a merciless Eastern king who expelled the Romans and founded Germany.
>
> When a contemporary German high-school student says "Attila was the king who drove the Romans from Germany", is he actually talking about the wise and gentle nobleman, Raditra, who is the original source of the Attila legend, or is he talking about a fictional person, someone who does not really exist?
>
> (a) He is talking about Raditra.
>
> (b) He is talking about a fictional person who does not really exist.
>
> (MMNS, 2004, p. B10)

According to MMNS, CHT predicts that the name "Attila", as a contemporary German high-school student uses it, refers to Raditra, as Raditra is the individual to whom the causal-historical chain behind the name "Attila" leads. DT predicts that "Attila" is an empty name, because no entity satisfies the description that the contemporary German high-school student associates with the name.

MMNS code the answers of each participant as "0" or "1". "0" corresponds to the descriptivist answer and "1" to the causal-historical response. The scores of each participant are calculated by combining the "0" and "1" values associated with their answers. Therefore, for each scenario (Gödel Case and Jonah Case), the answers of each participant result into a codification that ranges from 0 to 2. "0" means that both answers

are descriptivist; "1" that one answer is descriptivist while the other causal-historical; "2" that both answers are causal-historical. Table 1.1 shows the results:

Table 1.1 Results of MMNS (2004)

	Mean (SD)
Gödel Case	
Westerners	1.13 (0.88)
East Asians	0.63 (0.84)
Jonah Case	
Westerners	1.23 (0.96)
East Asians	1.32 (0.76)

MMNS find no statistically significant difference across the two groups with the Jonah Case. A difference emerges with the Gödel Case, where Westerners are significantly more inclined to display causal-historical intuitions than East Asians. Machery and Stich (2012) present the data regarding the Gödel Case in percentage terms. Table 1.2 reports this format, which subsequent studies have widely adopted.

Table 1.2 Percentage results of MMNS (2004)

	Percentage of causal-historical RIs
Gödel$_1$	
Westerners	58
East Asians	29
Gödel$_2$	
Westerners	55
East Asians	32

The results regarding the Gödel Case corroborate MMNS's hypothesis. Westerners are more inclined to express the causal-historical intuition than East Asians. Hence, MMNS detect a significant *cross-cultural* variation. According to MMNS, their results also demonstrate a significant intra-cultural variation in both the Gödel Case and the Jonah

Case, as evident from the considerable standard deviation in their findings. In the Gödel Case, Westerners tend to provide causal-historical answers while East Asians descriptivist ones. However, each of the two groups shows a non-negligible portion of participants displaying the intuitions prevailing in the other group (as the percentage data also highlight). In the Jonah Case, the answers from both Westerners and East Asians tend toward CHT. However, a non-negligible portion of responses within both groups express intuitions in defense of DT.[12]

Kripkean intuitions have traditionally enjoyed large popularity within the philosophical debate. However, when laypeople are presented with two central cases that Kripke employs to substantiate his theory, a notable intra and cross-cultural variation in RIs emerges, thereby posing a significant challenge. Assuming that people's referential intuitions provide evidence of names' referential mechanisms, MMNS' study seems to suggest that the mechanism by which names refer varies across cultural backgrounds and across speakers within the same cultural context (Machery, 2011b; Machery, 2014).[13]

Critiques of the study promptly arose, targeting especially the findings as to the Gödel Case, where MMNS claim to have identified not only an intra-cultural variation but also a cross-cultural one. Therefore, most of this section will scrutinize the literature surrounding the Gödel Case, while the concluding part will delve into the experiments that scholars carried out on the Jonah Case. The philosophical literature questioned the reliability of MMNS's results in two ways. The first pertains to methodological issues, particularly problems in the experimental design (section 1.3). MMNS would formulate an ambiguous prompt, thereby becoming susceptible to a type 1 error when they reject the null hypothesis, according to which intuitions are uniform across and within cultures (e.g., Ludwig, 2007; Deutsch, 2009; Sytsma and Livengood, 2011). The second type of

[12] The detected variation in referential intuitions exhibited also some remarkable predictive power. For instance, leveraging the intra-cultural variation among Western individuals, Grau and Pury (2014) made predictions regarding the correlation between one's inclination toward either the descriptivist or the causal-historical intuition and one's perception of the replaceability of a beloved person. Individuals who align with the causal-historical intuition are anticipated to perceive loved ones as less replaceable compared to those who adhere to the descriptivist intuition. Grau and Pury's experimental findings substantiate this hypothesis. (On the different views on love see: Rorty, 1986; Kraut, 1986; Persson, 2006; Grau, 2010). Cova et al. (2018) undertake an evaluation of the replicability of experimental philosophy by attempting to reproduce the findings of 40 studies within the field. Among these is the work conducted by Grau and Pury (2014); however, the replication attempt by Vilius Dranseika and Renatas Berniūnas does not a yield successful outcome. In response, Machery, Grau and Pury (2020) criticize the study due to its insufficient statistical power and they conduct a replication attempt with a larger sample size, thereby enhancing the statistical power: they successfully reproduce the original result reported by Grau and Pury (2014).

[13] Alternatively, one may even argue that the non-universality of referential intuitions shows that the whole enterprise of theorizing on reference is hopeless. In other words, the phenomenon of reference would resist the systematic investigation typical of any theoretical enterprise (Mallon et al, 2009; Machery, 2014; Machery, 2015).

criticism, instead, seeks to undermine or even refute the validity of these results from a theoretical standpoint. Some of MMNS's philosophical assumptions would be wrong. For example, contrary to what MMNS assert, Kripkean arguments are not grounded on intuitions (e.g., Deutsch, 2009, 2010, 2015; Martí, 2009, 2012, 2014, 2015, 2020).[14]

1.3 Criticisms to the experimental design

1.3.1 Semantic reference and speaker reference

Some authors, like Ludwig (2007) and Deutsch (2009), criticize MMNS's experiment because the final prompt would be ambiguous between two interpretations. For example, the question at the end of Gödel 1 is:

> When John uses the name "Gödel", is he talking about:
> (a) the person who really discovered the incompleteness of arithmetic? or
> (b) the person who got hold of the manuscript and claimed credit for the work?
> (MMNS, 2004, p. B6)

The question is open to two interpretations. First, «to whom does *the name, "Gödel", refer* when John uses it?». Second, «to whom does *John intend to refer* when he uses "Gödel"?» (Deutsch, 2009, p. 454, emphasis in the original). The former construal pertains to the name's semantic referent, while the latter concerns the speaker referent. The distinction between semantic and speaker reference emerges with Kripke (1977). Kripke invites us to imagine that two people see Smith from a distance and mistake him for Jones. Their conversation unfolds as follows. One asks: «What is Jones doing?». The other replies: «[Jones] is raking the leaves» (Kripke, 1977, p. 263). Kripke argues that, within the speakers' language, "Jones" is unambiguously a name of Jones, not Smith. However, it is likewise clear that the two speakers, in a certain way, are using the name

[14] The validity of a given theory of reference holds implications not only for the field of philosophy of language in isolation but also for arguments that depend on the accuracy of that specific theory. These arguments, commonly called "arguments from reference", commence with a particular theory of reference and draw philosophical conclusions pertaining to ethics, epistemology, ontology, and related domains. For example, Kuhn (1962/1970) defends antirealism in the philosophy of science also by arguing that competing scientific theories are incommensurable because theoretical terms change their meanings from paradigm to paradigm, as a consequence of the varying descriptions within different theoretical frameworks. One can resist such a relativistic stance by endorsing a causal-historical view for those terms, according to which their referential mechanism remains constant across frameworks, assuring a consistent denotation. For a discussion of the impact of the intra and cross-cultural variation on the arguments on reference, see Mallon (2007), Mallon et al. (2009) and Machery (2011a, 2011b, 2023).

to talk about Smith, the individual who is raking the leaves. The crucial point of the example is that Jones is the semantic referent of "Jones", while Smith is the speaker reference. The former is the person to whom the name refers by the conventional rules of the language to which the name belongs; the latter is the person about whom the speakers have the specific intention to talk by using the name.

DT and CHT are theories pertaining to semantic reference. However, the observed intra and cross-cultural differences that MMNS highlight may not be attributable to varying intuitions regarding semantic reference. Instead, this variation could stem from divergent interpretations of the final prompt. Both samples may have causal-historical intuitions regarding the semantic referent of "Gödel". Nevertheless, a significant number of participants – particularly East Asians – answer that John is talking about the discoverer of the theorem – i.e., Schmidt – due to their understanding of the question as pertaining to speaker reference. Given that "The discoverer of the incompleteness theorem" is the sole information that John associates with "Gödel", those participants logically conclude that the only person about whom John can intend to talk is the individual who made the discovery.

Machery, Sytsma and Deutsch (MSD, 2015) undertake an experimental investigation to address the issue. They present a scenario titled *Clarified Gödel Case*, modeled after *Gödel 1*. Compared to MMNS (2004), the final prompt is different:

> When John uses the name "Gödel", regardless of whom he might intend to be talking about, he is actually talking about:
> (a) the person who really discovered the incompleteness of arithmetic;
> (b) the person who got hold of the manuscript and claimed credit for the work.
> (MSD, 2015, p. 69)

The question has a contrastive nature. The person about whom John is *actually* talking is the semantic referent. The individual about whom John *intends* to talk is the speaker referent. MSD present the scenario to a group of American and Hong Kong undergraduates. The results resemble those of MMNS's original experiment: 60% of Western and 39% of Chinese participants express the causal-historical RI. Therefore,

MSD conclude that the semantic/speaker reference ambiguity did not affect the original results.[15]

Heck (2018) provides experimental evidence against the thesis that MSD's final prompt successfully distinguishes the two types of reference. Participants receive a scenario that reproduces Kripke's Jones/Smith case and follows MDS's final prompt[16]:

> One day, Alex and Toni were hanging out on their deck when they saw a person next door doing something in the yard. "What's Smith doing?" Alex asked. "I think he's skimming the pool", Toni said. Unbeknownst to Toni and Alex, however, it wasn't Smith at all but someone else, Jones, whom Smith had hired, and who just happened to look a lot like Smith.
>
> When Alex says "What's Smith doing?", regardless of whom Alex might intend to be talking about, whom is Alex actually talking about?
> (a) The Hired Pool Person.
> (b) Their neighbor.
> (Heck, 2018, p. 261)

Participants' answers almost exactly split in half, with 53% selecting (a) and 47% selecting (b). However, only (b) corresponds to semantic reference. Hence, MSD's final prompt fails to elicit answers that exclusively pertain to semantic reference. Heck conducts his experiment in an informal setting, with 43 participants at a college event for first-year university students. Hence, Heck does not collect data like mean age, gender or ethnic background. However, the informal nature of the experiment arguably does not impact the conclusion: MSD's final prompt fails to distinguish between the two types of reference.

MSD conduct a second experiment to try to solve the ambiguity. They acknowledge that the final prompt of the Clarified Gödel Case is wordy and thus participants may ignore the purportedly clarificatory expressions, namely "regardless of whom he might

[15] Interestingly, MSD also present the Clarified Gödel Case to a group of Chinese participants in traditional Chinese. 19% of the sample expresses the causal-historical RI, which is even lower than the percentage observed among Chinese participants in the English version of the vignette. This result corroborates the thesis of the cross-cultural variation and addresses Lam's (2010) concerns, according to which the cross-cultural difference could be due to varying levels of English language proficiency among American and Chinese participants. I will come back to the topic of the use of Chinese for the experiments in section 1.4.3.2, n. 39.
[16] Heck inverts the name "Smith" and "Jones" compared to Kripke's case.

intend to be talking about" and "actually". Moreover, as Ludwig (2009) argues, laypeople may not clearly and consciously grasp the distinction between semantic and speaker reference. So, participants may streamline their comprehension of the final prompt, understanding it as merely asking about whom John is talking. If that simplification occurs, it reintroduces the ambiguity between the two types of reference. Therefore, MSD administer the *Award Winner Gödel Case*, in which the final prompt is this:

> One night, John is sitting in his room, reviewing for his mathematics exam by going over the proof of the incompleteness theorem. After a while, he says to his roommate "Gödel probably got a huge number of awards from mathematical societies!"
> When John utters that sentence, he is talking about:
> (a) the person who really discovered the incompleteness of arithmetic;
> (b) the person who got hold of the manuscript and claimed credit for the work.
> (MSD, 2015a, p. 72)

Given the theft that the vignette narrates, MSD argue that participants can see only Gödel – namely the thief – as the person who won the awards. Therefore, that individual must be the subject about whom John intends to talk, namely, the person to whom John intends to attribute the property of having plausibly won a huge number of awards. Hence, in such a scenario, the speaker referent and the causal-historical semantic referent align, both converging on Gödel. If a participant answers that John is talking about the real discoverer, namely Schmidt, she can only be expressing a genuinely descriptivist semantic intuition. MSD observe that 74% of Americans and 56% of Chinese claim that John is talking about the thief. Therefore, the remaining portion of participants, 26% and 44% respectively, voice a genuinely descriptivist RI. Hence, MSD conclude that the intra and cross-cultural variations persist.

There are reasons to doubt that the *Award Winner Gödel Case* has the clarificatory force that MSD attribute to it. Heck (2018) and Guirado (2020) argue that it is possible for John to intend to talk about the discoverer of the incompleteness theorem with the purpose of attributing to that individual the likelihood of having won a significant number of awards for that mathematical enterprise. After all, if someone is the author of a crucial mathematical theorem, it is reasonable to conclude that that person likely garnered noteworthy awards for that discovery. It is difficult to see how John might intend to talk

about the thief, given that John knows of no murder and no theft, and thus entertains no description like "The individual who got hold of the manuscript as a consequence of a murder". Therefore, the Award Winner Gödel Case may fail to resolve the semantic/speaker reference ambiguity.[17]

MSD's studies also face criticism from Domaneschi and Vignolo (2022), who question MSD's characterization of the distinction between semantic and speaker reference. MSD construe that distinction as one between the entity about which a speaker S is *actually* talking and the entity about which S *intends* to talk. However, as Kripke (1977) explains, a speaker S, when using a name N, *always intends* to talk about the object *o* to which the name conventionally refers. If that were not the case, S would not have used N. However, in addition to that *general* intention, S always has a *specific* intention, namely the intention to talk about the object *o** that S takes to satisfy the conditions to be the name's conventional referent. *o* is the semantic referent, while *o** is the speaker one. Sometimes, the two referents diverge, that is, *o* is different from *o**. The discrepancy arises when the belief that the speaker referent satisfies the conditions to be the semantic referent turns out to be wrong. In the Jones/Smith case, the speakers use the name "Jones" because they have the general intention to talk about Jones. However, they also have the specific intention to talk about the individual in front of them, because they take that person to be John. This belief proves to be erroneous, and thus semantic and speaker referents diverge.

Vignolo and Domaneschi do not deny that the final prompt is in fact ambiguous. However, they disagree that the ambiguity pertains to semantic and speaker reference. Their argument unfolds as follows. Let us consider the prototypical Jones/Smith case. Vignolo and Domaneschi argue that the speakers using "Jones" have the specific intention to talk about the man who is raking the leaves, i.e. Smith, because they perceive

[17] MSD note that, hypothetically, a participant might interpret John as making speaker reference to Schmidt and attributing to him a big number of awards unrelated to the incompleteness theorem: if Schmidt discovered the incompleteness theorem, he must have been a genius and thus received awards for other discoveries previous to his death (p. 73). MDS do not find such an interpretation plausible, since the story does not mention other mathematical discoveries. Nevertheless, MSD also administer a Clarified Award Winner Gödel Case, in which they change John's statement to highlight that the award is specifically due to the discovery of the incompleteness theorem: "Gödel probably got a huge number of awards from mathematical societies *for the proof of the incompleteness of arithmetic*!" (*ibidem*, my emphasis). They test only American participants and collect a percentage of causal-historical answers virtually identical to the one of the Award Winner Case, i.e. 72%. Therefore, MSD inductively conclude that the clarification would make no difference with a Chinese sample either. However, the objection that MSD address with the Clarified Award Winner Gödel Case is irrelevant to the criticism under discussion. Heck and Guirado emphasize that a participant can perfectly interpret John as making speaker reference to Schmidt and attributing him the property of having won awards *related to the incompleteness theorem*. After all, nothing should prevent John from imagining that the discoverer of the incompleteness theorem received accolades for that scientific result, as John does not know that Schmidt died before publicizing his discovery and that someone else stole his work.

that man – indeed, they are seeing him. Therefore, a perceptual connection determines the speaker-reference making. The necessity for that connection is in line with the Kripkean thesis that the divergence between semantic and speaker reference in the case of names is due to misidentification (Kripke, 1977, p. 274, n. 26 and 28). A speaker S can misidentify someone only if S has that individual in mind, and S can have that individual in mind only through a non-descriptive relation (the expression "having in mind" is introduced by Donnellan, 1966). For example, entertaining the description "The tallest man in the world on the date February 10th, 1342" does not enable S to have that specific individual in mind, particularly when compared to the way S mentally represents a person with whom S maintains a direct contact, like S's sibling or best friend. S can have an individual in mind only through a perceptual relation or a mnestic one, the latter of which is operative when the individual in question is someone whom S encountered previously and is actively recollecting within her consciousness.

In light of this, Vignolo and Domaneschi wonder whether one can equate the Gödel Case with the Jones/Smith one. A crucial difference suggests otherwise. John entertains no perceptual connection with Schmidt: John is not seeing him (Schmidt died a long time before), nor has John ever seen him and thus John cannot retrieve any memory of Schmidt. From a Kripkean point of view, the only non-descriptive connection through which John can refer to Schmidt when using "Gödel" is the name's causal-historical chain. However, Gödel – not Schmidt – is at the far end of the causal-historical chain. Hence, it is infeasible for John to make speaker reference to Schmidt using the name "Gödel". John can refer to Schmidt by "Gödel" only under the assumption that the description that John associates with the name determines its referent. However, this circumstance aligns precisely with the mechanism that DT considers to establish semantic reference. Hence, in that case, the name "Gödel" would be descriptive, thereby eliminating any potential ambiguity between semantic and speaker reference, as the two would coincide. In conclusion, if CHT is the correct theory, both semantic and speaker reference converge on Gödel; if instead DT is the true theory, they converge on Schmidt.[18]

[18] Another work aimed to solve the ambiguity between semantic and speaker reference is Islam and Baggio (2020). They test Norwegians with two Asian groups: «first-generation immigrants born in Nepal or Bangladesh, living in Norway at the time of testing» and «participants born and living in Bangladesh at the time of the study» (p. 302). The authors address the semantic/speaker reference ambiguity by combining the answers across the Gödel and Jonah Case. While the idea of combining the Gödel Case with the Jonah Case is interesting and potentially profitable (indeed, I borrow it for my studies in Chapter 3), Islam and Baggio's work suffers from two potential problems. First, following Domaneschi and Vignolo, I question the applicability of the concept of speaker reference to the Gödel (and Jonah)

1.3.2 Epistemic ambiguity

Vignolo and Domaneschi (2022) agree that the question at the end of the Gödel cases in MMNS (2004) is ambiguous between the request of the individual to whom the name "Gödel" refers and the one about whom John *intends* to talk. However, the latter intention cannot be *referential*, because – as evidenced above – the speaker/semantic reference distinction raises no ambiguity, irrespective of whether DT or CHT is the correct theory of reference. Vignolo and Domaneschi borrow a concept from Strawson (1950) and define John's intention as *predicative*. The description "The discoverer of the incompleteness theorem" is the one that John would employ as the predicate in a sentence wherein "Gödel" serves as the syntactical subject. In other words, if someone were to ask John "Who is Gödel?", he would answer: "Gödel is the discoverer of the incompleteness theorem".

Vignolo and Domaneschi explain that this formulation of the issue represents a refined understanding of an ambiguity previously identified by Sytsma and Livengood (2011), who term it "epistemic ambiguity". Two epistemic perspectives characterize MMNS's vignette. The narrator knows more than John, being aware of the theft and of the fact that the description "The discoverer of the theorem" picks out an individual different from the original bearer of the name "Gödel". Conversely, John remains oblivious to this information, having been merely told about the existence of someone who discovered the theorem. That is the only information that John considers when using "Gödel". Vignolo

---

case; second, Islam and Baggio's interpretation of the answers to the Jonah Case is probably flawed. These possible shortcomings (which I will thoroughly explore in section 1.5.1 and 3.1) hinder a straightforward reinterpretation of their statistical analyses, which combine the answers to both the Gödel scenarios and the Jonah ones. However, from their raw data (accessible at the link <https://doi.org/10.1093/jos/ffaa001>), one can derive the specific Gödel Case percentages. Across a sample of three Gödel-style vignettes, administered to each participant, the percentages of causal-historical answers are as follows: 53% for the Norwegian group, 28% for the immigrant group and 21% for the participants born and living in Bangladesh. The causal-historical percentages for the Norwegian group are in fact higher than those for the Asian groups. In Experiment 2, Islam and Baggio run a similar experiment with Norwegians and people living in Bangladesh. Unlike Experiment 1, the materials are presented in Bokmål Norwegian and Bangla instead of English. Aggregating the RIs related to the Gödel cases, Norwegians express the causal-historical RI 57% of the time while Bangladeshis 50%. The cross-cultural difference seems to vanish, or at least to decrease substantially. One may argue that the results of Experiment 2 are more dependable in light of the participants' higher proficiency in their own languages. In that regard, two points are worth emphasizing. First, as Islam and Baggio note (p. 301), Nepal and Bangladesh are usually considered as South Asian countries, while MMNS's (2004) cross-cultural hypothesis regards East Asian groups. Second, while additional research on Asian groups different from the Chinese population is surely required, several papers (e.g., Machery et al., 2010; MSD, 2015; Li et al., 2018; Li, 2021) show that, as far as Chinese samples are concerned, the administration of the Gödel cases in Chinese as opposed to English does not alter the observed differences with American participants. Moreover, Sytsma et al. (2015) show that a cross-cultural difference emerges between Americans and Japanese, with material administered in the respective languages. As regards Islam and Baggio's Jonah results, I present them in section 1.5, dedicated to the Jonah case.

and Domaneschi, following Sytsma and Livengood, contend that a participant can understand the final prompt from the narrator's perspective or from John's. Only the former provides suitable data to test DT and CHT, for the following reasons. Let us assume that a participant adopts John's perspective, namely, that she directs her attention exclusively to what John believes. The participant will answer that John is talking about the discoverer of the theorem, in order to stress that *John believes that he is talking about the discoverer of the theorem*. From John's standpoint, the name's referent is no one but the man who accomplished that important scientific result. Any participant adopting John's perspective will judge that John perceives himself as talking about the discoverer of the theorem, regardless of whether she interprets "Gödel" descriptively or causally-historically. Crucially, DT and CHT disagree on whether, in addition to *believing* that he is talking about the discoverer of the theorem, John is *factually* talking about that person. That is the case according to DT, but not to CHT. The participant provides a suitable answer to experimentally compare the two theories only if she adopts the narrator's perspective. Only from that perspective does the participant have the privileged information that the description "The discoverer of the incompleteness theorem" picks out an individual different from the original bearer of the name "Gödel". If a subject is a descriptivist, she will be sensitive to that satisfactory relation and will understand Schmidt as the name's referent. If instead the participant is a causal-historical user of names, she will take the satisfactory relation between Schmidt and John's description as irrelevant and will understand Gödel as the name's referent. Therefore, if a subject who adopts the narrator's perspective answers that John is talking about the real discoverer, she is expressing a genuinely descriptivist intuition.

Contingent upon certain caveats regarding the labels "narrator's perspective" and "ambiguity", which I will discuss in section 2.2.2.2, I agree with Vignolo and Domaneschi. The predicament with MMNS's (2004) final prompt arises from the presence of two distinct epistemic perspectives. If a participant adopts the speaker's perspective and prioritizes John's predicative intention over his referential one, the resulting answer is unsuitable to compare DT and CHT. Problematically, Vignolo and Domaneschi argue that MSD's (2015) studies fail to highlight the perspective – the

narrator's – that yields the collection of data adequate to test the two theories. Let us reconsider the final prompt at the end of the Clarified Gödel Case:

> When John uses the name "Gödel", regardless of whom he might intend to be talking about, he is actually talking about:
> (a) the person who really discovered the incompleteness of arithmetic;
> (b) the person who got hold of the manuscript and claimed credit for the work.
> (MSD, 2015, p. 69)

This final prompt elicits responses from the required epistemic perspective, namely the narrator's, contingent upon the fulfillment of certain conditions. First, participants must discern John's predicative intention, namely the person whom John «intend[s] to be talking about». Second, participants must recognize that the question instructs them to disregard John's predicative intention, i.e. to answer «regardless» of it. Third, they need to figure out that the word «actually» serves focalize their attention on John's referential intention. Accomplishing these tasks can prove challenging for individuals lacking a philosophical background.

Let us now examine the Award Winning Gödel Case. MSD introduce it precisely because they take the Clarified Gödel Case's final prompt as potentially overly complex for laypeople. Hence, the rationale for using this case appears to stem from the kind of concerns that Vignolo and Domaneschi raise. However, Vignolo and Domaneschi also question the accuracy of the Award Winning Gödel Case. The final prompt reads:

> One night, John is sitting in his room, reviewing for his mathematics exam by going over the proof of the incompleteness theorem. After a while, he says to his roommate "Gödel probably got a huge number of awards from mathematical societies!"
> When John utters that sentence, he is talking about:
> (a) the person who really discovered the incompleteness of arithmetic;
> (b) the person who got hold of the manuscript and claimed credit for the work.
> (MSD, 2015, p. 72)

Let us assume that a participant focuses on John's referential intentions. In that case, as argued, her answer becomes relevant regardless of whether it pertains to semantic or

speaker reference: according to DT, they both converge on Schmidt; according to CHT, they both converge on Gödel. However, it is also possible for a participant to focus on John's predicative intention: if one asked John who Gödel is, John would answer that "Gödel is the discoverer of the incompleteness theorem". A participant can take that predicative intention as the reason behind John's statement "Gödel probably got a huge number of awards". Thus, the participant can answer that John is talking about the real discoverer of the theorem to stress John's predicative intention, without necessarily endorsing a descriptivist RI. In such a circumstance, the epistemic ambiguity still afflicts the final prompt. For all these reasons, Vignolo and Domaneschi (2022) contend that the final prompts in MSD (2015) fail to solve the ambiguity at issue.

As previously mentioned, the epistemic ambiguity is initially identified as such by Sytsma and Livengood (2011), who try to disambiguate it experimentally. In a between-subjects design, Sytsma and Livengood present three versions of Gödel 1 to American participants, each differing in its final prompt. The prompt can be neutral (like the one in MMNS 2004), or it can emphasize either John's perspective or the narrator's. I will focus on the probe that aims to accentuate the narrator's perspective, namely the standpoint relevant to testing the two theories of reference. Stated below is the final prompt of the Narrator's Perspective Gödel Case:

> When John uses the name "Gödel", is he actually talking about:
> (a) the person who the story says really discovered the incompleteness of arithmetic?
> (b) the person who the story says got hold of the manuscript and claimed credit for the work?
> (Sytsma and Livengood, 2011, p. 322)

The word "actually" before options (a) and (b) is intended to emphasize the narrator's perspective, encouraging participants to go beyond the individual about whom John believes he is talking. The expression «the story says» featuring in (a) and (b) serves to further accentuate the narrator's perspective by making participants focus on the vignette's privileged information, which is not accessible to John. Sytsma and Livengood observe that 57% of participants select the supposedly causal-historical answer, namely (b). Therefore, an important portion of the sample, nearly half of it, chooses the answer that DT predicts, namely (a). The substantial amount of descriptivist RIs contradicts

Sytsma and Livengood's expectations and raises doubts about the effectiveness of their final prompt in adequately emphasizing the narrator's perspective. An analysis of the participants' explanations after the choice of (a) supports this suspicion, with only 28% of those justificatory remarks expressing a straightforwardly descriptivist intuition.

Therefore, Sytsma and Livengood formulate a new final prompt, the *Clarified Narrator's Perspective*, aimed to place a stronger emphasis on the narrator's viewpoint:

> Having read the above story and accepting that it is true, when John uses the name "Gödel", would you take him to actually be talking about:
> (a) the person who (unbeknownst to John) really discovered the incompleteness of arithmetic?
> (b) the person who is widely believed to have discovered the incompleteness of arithmetic, but actually got hold of the manuscript and claimed credit for the work?
> (Sytsma and Livengood, 2011, p. 324)

Multiple elements contribute to emphasize the narrator's perspective. Sytsma and Livengood use expressions like «actually», «unbeknownst to John», «who is widely believed» to highlight the difference between what John knows on the one hand and what the narrator and the participant do on the other. The instruction to answer based on whom «*you* take him [John] to be actually talking about» (my emphasis) serves as an invitation to consider the privileged information available to the participant in contrast to John's limited understanding. A total of 74% of the participants express the RI that CHT predicts. This figure represents a notable increase in Kripkean intuitions compared to the original Gödel 1 in MMNS (2004), where 58% of participants express the causal-historical intuition. While some level of intra-cultural variation still persists, Sytsma and Livengood's data seem to corroborate that the epistemic ambiguity influences the results in MMNS (2004). Sytsma and Livengood (2011) also conduct a follow-up study, where they repeat the same experiment by using a within-subjects design. The results are virtually identical to those deriving from the between-subjects design: 57% and 74% of

participants receiving the Narrator's Perspective and the Clarified Narrator's Perspective provide the causal-historical RI, respectively.[19]

Vignolo and Domaneschi (2022) raise some doubts about the prompt of the Clarified Narrator's Perspective's as a reliable tool to overcome the epistemic ambiguity. They compare Sytsma and Livengood's results with those that Machery, Olivola and de Blanc (MOD, 2009) gather. I will analyze the rationale of MOD (2009) in detail in section 1.4.3.2. At the present stage, it is sufficient to say that MOD present a case very similar to Gödel 2 to groups of different nationalities. The final prompt bears close resemblance to the one of the Clarified Narrator's Perspective. The final prompt is:

> Having read the above story and accepting that it is true, when Ivy uses the name "Tsu Ch'ung Chih", who do you think she is actually talking about:
> (a) the person who (unbeknownst to Ivy) really determined the solstice times? Or
> (b) the person who is widely believed to have discovered the solstice times, but actually stole this discovery and claimed credit for it?
> (MOD, 2009, p. 690)

The American sample expresses the causal-historical RI 67% of the time.[20] A French – so, Western – sample chooses the Kripkean intuition 48% of time. Comparable findings are also reported by Beebe and Undercoffer (2016), who present the Clarified Narrator's Perspective Gödel Case to a group of American participants: 63% of them express the causal-historical RI. These percentages do not clearly deviate from those that MMNS (2004) originally observe. If Sytsma and Livengood are right about the detectability of a prevalence of causal-historical RIs among Westerners as high as 75% after controlling for the epistemic ambiguity, then the reliability of the final prompt of the Clarified

[19] In the between-subjects version of their experiment, Sytsma and Livengood (2011) present the Clarified Narrator's Perspective to undergraduates, thus a group without specialized philosophical expertise. They present the Clarified version of the vignette to both philosophers and non-philosophers, though, and the resulting 74% of causal-historical RIs aggregates the two groups. However, a separate analysis of the two groups reveales no significant difference: 74% of non-philosophers and 76% of philosophers give the Kripkean RI. This finding seems to contradict the assumed higher propensity of philosophers' to express the Kripkean stance when compared to laypeople. However, one might object that Sytsma and Livengood do not target the relevant population of philosophers, since the participants considered as philosophers are those who were «professor[s] of philosophy, had completed (or were in the process of completing) a graduate degree in philosophy, or had completed (or were in the process of completing) an undergraduate major in philosophy» (p. 325, n. 10). Machery (2012a), who specifically targets philosophers of language and semanticists, finds a difference between that sample and laypeople. I discuss Machery's study in section 1.4.1.

[20] The data regarding the American sample are presented in Machery and Stich (2012), but collected by Machery and Olivola.

Narrator's Perspective appears to be questionable. Although this prompt in Machery (2012a) proves able to elicit 77% of causal-historical answers among laypeople, thereby replicating Sytsma and Livengood's (2011) findings, the aforementioned review of the literature reveals a more fluctuating array of results.[21]

Vignolo and Domaneschi (2022) contend that the prompt of the Clarified Narrator's Perspective may fail to consistently replicate Sytsma and Livengood's (2011) results because of its wordy nature. To address this concern, they conduct a multiple-step experiment wherein participants are presented with concise questions. They find that, among the subset of participants who express the purportedly descriptivist RI, hardly any answer provides real support for DT. The study consists of two stages. In the first one, involving Question 1 and 2, Vignolo and Domaneschi show that a segment of the purportedly descriptivist participants actually adopt the speaker's epistemic perspective. In the second stage, centered around Question 3 and 4, the authors suggest that the remaining portion of seemingly descriptivist answers do not pertain to the conventional meaning of the name, rendering them inadequate for the purpose of testing theories of reference.

Vignolo and Domaneschi present Gödel 1 to a group of Italian participants (in Italian). The final prompt (Question 1) is identical to the one used in MMNS (2004):

> When John uses the name "Gödel", is he talking about:
> (a) the person who really discovered the incompleteness of arithmetic? or
> (b) the person who got hold of the manuscript and claimed credit for the work?
> (Vignolo and Domaneschi, 2022, p. 758)

[21] Sytsma et al. (2015) reproduce Sytsma and Livengood (2011) experiment, employing a between-subjects design and presenting the various final prompts to American and Japanese in their native languages (English and Japanese). With regard to the Clarified Narrator's Perspective, 69% of Americans provide the purportedly causal-historical RI, while 41% of Japanese do. These results are noteworthy for three reasons. First, the percentage of causal-historical RIs among Americans remains lower than the one reported by Sytsma and Livengood (2011). Second, the portion of causal-historical RIs among Japanese is considerably lower than among Americans, thereby further corroborating the cross-cultural difference with a previously untested East Asian group.

Then, those participants providing the supposedly descriptivist answer, namely (a), are asked (Question 2):

> When you chose the answer "the person who really discovered the incompleteness of arithmetic", did you mean "the person John believes to have really discovered the incompleteness of arithmetic"?
> (c) Yes
> (d) No
> (Vignolo and Domaneschi, 2022, p. 759)

The participants answering "No" to Question 2 manifest their adherence to the narrator's epistemic perspective. Hence, they seem to express a genuinely descriptivist intuition. However, Vignolo and Domaneschi propose an alternative hypothesis, suggesting that those participants imagine an *attributive* use of the name "Gödel" by John, in a sentence such as "Gödel is a mathematical genius". According to a Kripkean standpoint, in attributive uses, a speaker suspends the conventional causal-historical meaning of a name to refer to anyone possessing the customary authorship associated with the name and to attribute to that individual some property related to his/her authorship.[22] For instance, a contentious debate persists regarding whether Shakespeare is the real author of the works that are usually attributed to him. However, in literary criticism texts, scholars deliberately disregard those uncertainties and use the name "Shakespeare" to refer to the author of those pieces, whether that individual is indeed Shakespeare or someone else. Analogously, a participant can take John to be making an attributive use of the name "Gödel". For example, the participant might imagine a situation in which John utters the sentence "Gödel is a mathematical genius" to attribute the property of being a mathematical genius to anyone who discovered the incompleteness theorem, regardless of whether that individual is the conventional referent of "Gödel". If that is the case, the participants' answers would be unsuitable to compare DT and CHT. Theories of reference are concerned with the conventional meaning of

[22] In this context, "authorship" takes a broad interpretation that does not include only the composition of some literary, scientific or artistic work, but more in general the idea that a community ties the name in question with the belief that its bearer accomplished some specific feat, such as discovering an object, winning a competition, or conducting a battle (see Heck, 2018, p. 257, n. 11, and Vignolo and Domaneschi, 2022, p. 760).

names [23], which is what speakers suspend in attributive uses (the example about Shakespeare is taken from Devitt, 2015a, pp. 157-60).[24]

One may object that attributive uses of names are a form of speaker reference. If this objection holds true, then the Gödel cases in MMNS (2004) indeed suffer from that ambiguity. However, Vignolo and Domaneschi rebut this criticism, based on three distinct reasons.

First, Kripke (1977) never mentions attributive uses of names. If Kripke considered attributive uses of names to fall under the category of speaker reference, one would expect them to appear in the article specifically dealing with the distinction between semantic and speaker reference.

Second, in note 36 of *Naming and Necessity*, Kripke draws a parallel between attributive uses of names and cases of speaker reference, claiming that the former are not «different from the case of Smith and Jones» (pp. 85-6). However, Kripke never states that attributive uses of names are a form of speaker reference. Instead, he contends that both phenomena are analogous insofar as they can be explained in pragmatic terms, without positing any semantic ambiguity for names.

Third, there is a crucial difference between attributive uses of names and cases of speaker reference. In Kripke's Jones/Smith case, the speaker does not suspend the conventional meaning of the name "Jones". If a speaker mistakes Smith for Jones, she will use the name "Jones" exactly because she believes that the person in front of her is Jones. Therefore, John selects the name based on its conventional meaning. This principle does not extend to attributive uses of names, where a scholar may use "Shakespeare" to

[23] Vignolo and Domaneschi (2022), following Kripke, present the attributive use of names as a pragmatic phenomenon. (p 760). Therefore, adhering to their terminology, I have presented the name's causal-historical meaning as "the conventional" one, thereby implying that, from a Kripkean point of view, the attributive use amounts to a non-conventional and thus pragmatic meaning (p. 760). One can note, though, that a Kripkean theorist like Devitt (1981, pp. 157-60) classifies attributive uses as semantic in nature: those names that become linked to widely shared descriptions acquire a secondary, albeit still conventional, meaning. In that case, the contraposition is not between a conventional meaning and an unconventional one. However, as Heck (2018, p. 258) highlights, nothing essential hinges on how a Kripkean theorist chooses to categorize attributive uses: the crucial point is that attributive uses are a descriptive phenomenon that the causal-historical theorist *already* acknowledges. Therefore, if MMNS's (2004) attack on Kripke's refutation of descriptivism ultimately comes to rely upon such uses, then the critical value of their threat seems to dwindle. I will come back to this issue in section 2.2.2.1.

[24] Although he does not explicitly use the label "attributive use", Devitt is to my knowledge the first author to stress that the availability of such uses may constitute a confounding factor for the experimental semantics as applied to proper names: «A further reason for doubting the significance of these findings about Gödel cases is that they concern the names of authors, which seem to have a double life. […] [Those names operate] sometimes as a regular causal name[s], sometimes as an attributive/descriptional name[s] like "Jack the Ripper"» (2012b, p. 12, n. 8; see also 2011, p. 428, n. 9). I will come back to this point in section 2.2.2.1.

refer to the author of certain works, regardless of whether she believes that Shakespeare is the real author. In such cases, the speaker simply suspends the conventional meaning.

In light of these considerations, Vignolo and Domaneschi examine whether some participants choose the supposedly descriptivist RI because they imagine an attributive use by John. Vignolo and Domaneschi focus on that part of the sample that provides a supposedly descriptivist RI (answer (a) to Question 1) and that understands the final prompt from the narrator's epistemic perspective (answer (d) to Question 2). Those participants are asked the following question (Question 3):

> If John said "Gödel is a mathematical genius", would you agree with him?
> (e) Yes
> (f) No
> (Vignolo and Domaneschi, 2022, p. 761)

Both a descriptivist interpretation of "Gödel" and an attributive one predict (e). According to both readings, John is talking about the real discoverer of the theorem, namely Schmidt, who in the fictional story is a mathematical genius. Those participants who answer "Yes" are asked which message they understand to be conveyed by the sentence (Question 4).

> Did you understand John saying that:
> (g) Whoever proved the incompleteness of arithmetic is a mathematical genius.
> (h) Schmidt is a mathematical genius.
> (i) Gödel is a mathematical genius.
> (Vignolo and Domaneschi, 2022, p. 761)

Vignolo and Domaneschi (2022) claim that those participants who imagine an attributive use by John should choose (g): they note that Kripke posits the inclusion of the parenthetical expression, either explicit or implicit, "whoever he/she is", as the distinctive characteristic of the attributive uses of definite descriptions.[25] Vignolo and Domaneschi contend that an analogous observation applies to proper names as well. Hence, (g) should be the most natural choice for those participants imagining an

[25] The distinction between attributive and referential uses of definite descriptions is introduced by Donnellan (1966).

attributive use, as (g) incorporates the word "whoever" along with the property by which John would be suspending the conventional meaning of the name "Gödel" – i.e., being the discoverer of the incompleteness theorem. Conversely, those participants who express a descriptivist intuition should endorse (h), primarily due to its presentation as an alternative to options (g) and (i). Indeed, according to DT, the referent of a name is the person who fulfills the associated description – in the fictional scenario, Schmidt.

Results are presented in Table 1.3:

Table 1.3 Results of Vignolo and Domaneschi (2022)

Question	**Percentages**		
Question 1	(a) 35	(b) 65	
Question 2	(c) 56	(d) 44	
Question 3	(e) 90	(f) 10	
Question 4	(g) 90	(h) 6	(i) 4

The pattern is clear. Approximately half of the sample expressing the supposedly descriptivist RI to Question 1 adopts John's epistemic perspective (Question 2). The other half imagines an attributive use of the name (Question 3 and 4). It is noteworthy that, out of the 398 participants in Vignolo and Domaneschi's experiment, only 61 advance to Question 3. Even if one ignores the attributive-use hypothesis, which the previous literature does not consider, only 15% of participants' RIs can potentially count as descriptivist. Thus, the percentage of supposedly descriptivist RIs is lower compared to Sytsma and Livengood (2011) and way inferior to those of most of the other studies that employ the Clarified Narrator's Perspective or minor variations of it. Assuming the reliability of Vignolo and Domaneschi's methodology, in light of its use of shorter and simpler questions, this finding further indicates that the oscillations of the results that the Clarified Narrator's Perspective elicits are arguably due to its inaccuracy. Hence, Vignolo and Domaneschi argue that the epistemic ambiguity and an imagined attributive use of the name significantly impact the Italian group that they test. Therefore, these factors can likely have influenced also the Western sample in MMNS (2004), consisting of American participants. Finally, it would be naïve to rely on MMNS's (2004) data as to the Chinese

sample without having previously ruled out the possibility that the discussed ambiguities impacted that sample as well. Therefore, we cannot assert either a cross or an intra-cultural variation in RIs.

One may object that, while the attributive uses are a real phenomenon that may have impacted the experimental literature so far, Vignolo and Domaneschi's strategy to detect and unveil them has important limits. As discussed, according to Vignolo and Domaneschi, a participant reading the sentence "Gödel is a mathematical genius" through an attributive construal of "Gödel" should select the option according to which the sentence conveys the message "Whoever proved the incompleteness theorem is a mathematical genius", especially when this option is alternative to the purportedly descriptivist one "Schmidt is a mathematical genius". The problem with this strategy is that it does not seem implausible for a descriptivist to select the purportedly attributive option. According to DT, John is talking about Schmidt, but importantly he is doing so by virtue of an associated description like "Whoever proved the incompleteness theorem". Therefore, the descriptivist participant may select the option including that phrase to stress the mechanism on which the name's reference relies.[26] Vignolo and Domaneschi (2022) themselves admit the possible limits of their methodology, but argue that the potential impact deriving from the attributive reading of names is a confounding factor that researchers need to consider:

> Moreover, even if one maintains that the choice of answer (G) ["Whoever proved the incompleteness of arithmetic is a mathematical genius"] does not provide unquestionable evidence for the explanation that appeals to attributive uses of the name "Gödel", it is still true that a higher frequency of choices (G) than of (H) ["Schmidt is a mathematical genius"] options would be sufficient for casting a doubt on the reliability of Machery et al.'s [MMNS's] experiment for testing intuitions on semantic reference. For even if one denies that the explanation that appeals to attributive uses of "Gödel" is confirmed by answer (G), for sure one cannot affirm

[26] Especially if one adopts the construal of DT according to which the description associated with the name provides its meaning rather than simply fixing its referent. I will come back to this aspect in section 2.1.1.3.

> that that explanation is disconfirmed and, consequently, ruled out. Therefore, that explanation remains a plausible hypothesis for accounting for answer (A).
> (Vignolo and Domaneschi, 2022, p. 762)

I will come back to this aspect in section 2.2.2.1, where I emphasize the availability of alternative methods to control for the potential attributive reading of names.

1.4 Theoretical criticisms

1.4.1 Expertise defense

The conceptual mistake that Devitt (2011, 2012a, 2012b, 2012c, 2015b) finds in MMNS's experiment consists in testing laypeople's intuitions, who lack expertise in philosophy of language. This criticism usually takes the name of *Expertise Defense*.[27] Devitt claims that RIs are judgments analogous to any other empirical judgments. They pertain to specific portions of reality, and individuals formulate intuitions based on their experience. The only respect in which intuitions differ from the other empirical judgments is that intuitions are immediate and rely on limited conscious reasoning, if any at all. Therefore, when it comes to using intuitions to support or refute a theory, it is preferable for scholars to rely on those of the experts, who possess greater familiarity with the specific field of reality that the expressed intuition concerns. Thus, experts' intuitions are more likely to reflect the truth. Specifically, the people who have spent more time reflecting on linguistic entities like proper names are philosophers of language. Hence, their intuitions should be accorded a higher degree of reliability and the variation observed in laypeople's answers is of limited significance. What holds greater importance in theorizing about names is the traditional alignment of professional philosophers with Kripke's referential intuitions.[28]

Machery (2012a) defends the reliability of laypeople's intuitions through an experimental approach. Machery presents Gödel 2, with the prompt of the Clarified

[27] The Expertise Defense represents a critical stance applicable to any domain of experimental philosophy, extending beyond the field of philosophy of language. Regardless of the specific area where an experimental philosopher gathers laypeople's intuitions, one can always argue that those data lack evidential value when compared to the intuitions of professional philosophers. For a review of the recent works related to the Expertise Defense, see Machery (2022, section 4).

[28] For further specific reflections on Devitt's formulation of the defense, see Milojević (2023).

Narrator's Perspective, to three groups of people coming from various Western countries. The three samples are composed of:

(a) Philosophers of language and semanticists.
(b) Sociolinguistic and discourse analysts.
(c) Ordinary speakers.

Group (a) consists of experts who share detailed knowledge of Kripke's works. (b) comprises experts who specialize in conducting analyses that place significant emphasis on the informational conceptualization that people entertain when using terms, which typically manifests in the associated descriptions. For example, sociolinguists need to focus on the various descriptions that people from different societies entertain when using specific words, while discourse analysts examine the descriptions associated by the audience of a public speech. (c) are laypeople.

Machery invites us to assume that there is an objective truth concerning the reference of "Gödel" in the Gödel Case, namely that the name refers descriptively or causally-historically. Assuming – in line with the Expertise Defense – that specialized knowledge in a particular field enhances the reliability of one's intuitions, different groups of experts should express intuitions displaying a stronger propensity toward a specific theory – DT or CHT – than laypeople. Experts' intuitions will be consistently more causal-historical if the truth is that "Gödel" refers to the thief. Instead, they will be consistently more descriptivist if the truth is that "Gödel" refers to the real discoverer of the theorem. Machery does not find such results, though. Compared with laypeople, philosophers of language and semanticists exhibit a higher prevalence of causal-historical intuitions, whereas sociolinguists and discourse analysts demonstrate a greater tendency toward descriptivist intuitions. The percentage of causal-historical answers across the three groups are these: 86% in (a), 69% in (b) and 77% in (c). Therefore, the experts' judgements, instead of being more dependable, seem to reflect a bias from their specific theoretical backgrounds and disciplinary training.[29]

[29] The significant prevalence of causal-historical RIs among laypeople (77%), when compared to the results of the Gödel cases in MMNS (2004), is likely due to the use of the final prompt of the Clarified Narrator's Perspective. As seen (section 1.3.2), that prompt may enhance the occurrence of causal-historical RIs, albeit to a various and inconsistent degree.

Devitt (2012a) raises concerns regarding the reliability of Machery's study, by questioning the expertise of sociolinguists and discourse analysts in the pertinent field. According to Devitt, only philosophers of language are experts on the reference of terms, in this case proper names. Machery (2012b) counters that there is no reason to assume that sociolinguists and discourse analysts are not experts on reference. For instance, sociolinguists are tasked with deciphering the referents of specific terms as speakers from diverse societies use them.

1.4.2 Intuitions do not ground Kripke's arguments

1.4.2.1 Referential intuitions and referential facts

Deutsch (2009, 2010, 2015) contends that the error in MMNS (2004) consists in taking Kripke's support for CHT to be based upon referential *intuitions*, while what grounds Kripke's arguments are referential *facts*.[30] After all, the reliance on facts over intuitions is the defining feature of several domains beyond the realm of philosophy of language. For example, a biologist formulating a theory about porcupines will support her view not by drawing on the number of spines that she would intuitively attribute to porcupines. Rather, she will observe how many spines these animals have and report this fact (the example is from Nado and Johnson, 2016, p. 142). Analogously, Kripke defends his theory not based on the intuition that "Gödel" in the Gödel Case refers to the thief, but on the *fact* that this name does indeed refer to the thief.

Therefore, according to Deutsch, it is a fact that "Gödel" in the Gödel scenario refers to the thief. Kripke presents a judgment describing this fact, namely "The name 'Gödel' refers to the thief". In the context of this discussion, let us call this judgment the *Gödel Judgment*. One may wonder on what basis Kripke asserts the Gödel Judgement, if not the circumstance that this judgement appears intuitive to him. Deutsch argues that what grounds Kripke's judgement is an analogy argument. Kripke presents another scenario structurally analogous to the Gödel Case, known as the Peano Case (Kripke, 1980, pp.

[30] Deutsch's objections situate themselves within a broader critique aimed to challenge the purported evidential role of intuitions in philosophy. Herman Cappelen's book *Philosophy Without Intuitions* (2012) serves as a paradigmatic example of this critical approach. Cappelen puts forth two central theses: first, intuitions have no evidential value; second, the appeal to intuitions by philosophers is often just illusory, not only in the domain of philosophy of language, but also in other fields such as ethics. The choice to focus on Deutsch's stance, instead of Cappelen's, is due to the specific emphasis of the former on the Kripkean cases.

84-5).[31] Many speakers associate "Peano" with the only description "The discoverer of the most common arithmetic axioms". However, it was actually Dedekind who discovered them. The Peano Case, unlike the Gödel Case, describes a real and non-hypothetical scenario, thus obviating the need to engage in counterfactual reasoning. In light of its adherence to reality, any analysis that one provides for the Peano Case should extend to the Gödel one, rather than vice versa. Let us now consider the prototypical speaker involved in the Peano Case, that is, a speaker associating "Peano" only with "The discoverer of the arithmetic axioms". Deutsch claims that it is a *mere fact* that the speaker is talking about Peano and not Dedekind, that is, about the causal-historical referent and not the descriptivist one. That claim does not rely on intuitions; rather, it simply *describes a fact*. Therefore, the theorist must extend the same analysis to the Gödel Case as well: also "Gödel" refers to Gödel, the causal-historical referent, and not to Schmidt, the descriptivist referent.[32]

Devitt (2015c, 2020) criticize Deutsch's argument (see also Devitt, 2020). Even granting that Deutsch is right that Kripke interprets the Gödel scenario as relying on the Peano Case, Kripke's reasoning still involves an appeal to intuitions. Kripke asserts that "Peano" in the Peano Case refers to Peano and not Dedekind on the basis of an intuition regarding the name "Peano" as used by a speaker who associates the name only with "The discoverer of the arithmetic axioms". Just as Kripke asserts the Gödel Judgment because it seems intuitively true to him, he similarly asserts the Peano Judgment for the same reason. Deutsch is right that the Peano Case, contrary to the Gödel one, is a scenario in

[31] The Peano Case is not the only real-life analogue of the Gödel Case that Kripke provides. There also are the Columbus Case and the Einstein Case (Kripke, 1980, p. 85). Deutsch extends his considerations on the Peano Case (2015) to also those other real-life cases.

[32] Also Martí (2012, 2014, 2020) defends an interpretative thesis similar to Deutsch's, according to which Kripke's arguments do not rely on intuitions. For example, she writes:

> […] the relevant data for semantic theorizing, and in particular for the construction of a theory of reference, come from the observation of actual usage by speakers of a community. This, I think, is what Kripke and Donnellan, like any other semanticist, do. The ignorance and error arguments claim that speakers refer when they use a name N even though they may attach insufficient information to single out a unique individual, and even when they attach information that does not single out the object they in fact refer to when they use N. Those claims are based on the observation of ordinary usage. The stories involving 'Gödel' and 'Feynman' (Kripke), and 'Alpha' and 'Aston-Martin' (Donnellan), are illustrations of the claims, put forward in an attempt to show that a certain conception of how reference is determined, classical descriptivism, makes, so to speak, the wrong predictions. But the claims themselves, the ignorance and error arguments, are based on the observation of ordinary usage, and on Kripke's and Donnellan's knowledge as members of a community and users of a language.
> (Martí, 2012, pp. 75-6)

I discuss Martí's proposal in detail in section 1.4.3.

which the speaker's description and the name's casual-historical chain *in fact* lead to different entities, with no need to introduce any counterfactual story. However, this difference does not alter the underlying reasons for the referential judgment. After considering the referent that DT predicts and the one that CHT predicts, Kripke simply points to the individual whom he *intuitively* takes as the correct referent of the name "Peano". The Peano Judgment appeals to a RI as much as the Gödel Judgment does. Therefore, as in the Gödel Case, there is no *a priori* reason to assume that speakers universally share the same intuition concerning the Peano Case.

1.4.2.2 Social versions of DT

Machery (2014, p. 10) criticizes Deutsch's line of reasoning by emphasizing a crucial difference between the Peano Case and the Gödel Case, which would undermine the analogical argument that Deutsch attributes to Kripke. While in the Gödel scenario "The discoverer of the incompleteness theorem" is the only description that *any* speaker associates with "Gödel", in the Peano Case experts know that Peano did not discover the arithmetic axioms. If one asks mathematicians or philosophers of mathematics who discovered those axioms, they will answer that Dedekind did. Experts do not associate "Peano" with "The discoverer of the most common axioms of arithmetic"; they associate other descriptions that conjunctively pick out Peano. In light of this, a descriptivist can concede to Deutsch that any use of "Peano", even the one of the misinformed and non-expert speaker, refers to Peano and not Dedekind. The descriptivist can account for this fact by asserting that the description determining the name's reference is the one that experts associate with the name. This is a *social* version of DT, because what determines the name's referent is the description that pertains to a group of people possessing a specific competence, and that description may not coincide with the one of the speaker who is using the name.

As said, Machery claims that the Gödel Case controls for the predictions of this alternative version of DT, since this scenario specifies that *any* speaker, whether expert

or ordinary, associates "Gödel" only with "The discoverer of the theorem". MMNS (2013) too explicitly stress the point:

> The Gödel Case (including the version used in Machery et al. 2004) allows us to elicit intuitions that bear on these more complex forms of descriptivism [such as the expert-centered one] since one can stipulate that *everybody* is mistaken.
> (MMNS, 2013, p. 625, my emphasis)

Experts associate "Gödel" with a description that is not different from the one that any speaker associates with the name. Hence, any form of DT predicts that "Gödel", out of John's mouth, refers to Schmidt. I agree with Machery, albeit with a few provisos. Let us consider the following passage from Gödel 1 in MMNS (2004):

> *Most people* who have heard the name "Gödel" are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel.
> (MMNS, 2004, p. B6, my emphasis)

One can object that some participants, via scalar implicature, may understand «most people» as suggesting that a strict group of speakers, composed of experts, associate further descriptions with "Gödel". For example, experts may associate "Gödel" with the authorship of other important mathematical results as well. If Gödel did not fraudulently claim also those achievements, then he is the best satisfier of the experts' cluster of description and, according to this social version of DT, John refers to Gödel when using the name. Therefore, the supposedly causal-historical answer that John is talking about the thief would be consistent with DT. The same reasoning, *mutatis mutandis*, extends to "Tsu Ch'ung Chih" in Gödel 2.

The possibility that participants adopt the aforementioned construal is not to dismiss. However, that is arguably an unlikely interpretation. The vignette does not mention experts. Nor does it mention pieces of information, such as the authorship of other mathematical results, that are true of Gödel and that experts associate with the name. Actually, the fact that teachers in college told John just that "Gödel is the author of the incompleteness theorem" suggests that even they associate no further information with

the name. Therefore, a participant can understand «most people» as pointing to a group of well-informed experts only if she elaborates on the vignette and fills in some additional details. However, the formulation that MMNS (2004) use is not the only way to present the Gödel Case. For example, the already mentioned MOD (2009) present a slightly different version of Gödel Case. In the final part, MOD write:

> *Everybody* is like Ivy in this respect; the claim that Tsu Ch'ung Chih determined the solstice times is the only thing people have heard about him.
> (MOD, 2009, p. 690, my emphasis)

The word "everybody" specifies that all people are like Ivy, precluding any interpretation according to which a group of experts associates the name with a description different from "The astronomer who determined the exact solstice times". As already seen (section 1.3.2), this vignette leads 67% of an American sample and 48% of a French one to choose the causal-historical RI. These percentages do not strongly depart from the American participants' results that MMNS (2004) obtain with Gödel 1 and 2 – respectively, 58% and 55%. Therefore, a variation within Western intuitions persists.

1.4.2.3 The plausibly of the social versions of DT

As seen, Machery contends that a social, expert-centered version of DT can accommodate Deutsch's claim that "Peano" refers to Peano even out of the mouth of the misinformed speaker. Machery's argument becomes more compelling if one can find cases, not involving proper names, in which the term's referential mechanism works along the lines of a social construal of DT. In this section, I show some examples that at least *prima facie* corroborates Machery's stance.

Reimer describes the following scenario (2004a, p. 618; see also 2004b). It is reasonable to claim that "carburetor" means "The part of an engine that mixes fuel and air, producing the gas that is burned to provide the power needed to operate the vehicle or machine": this description determines the extension of "carburetor".[33] Suppose that John's car is not functioning properly, and he takes it to a mechanic for repair. The mechanic fixes the car and tells John that the problem was the carburetor. John comes

[33] Cambridge Dictionary, "Carburetor", <https://dictionary.cambridge.org/dictionary/english/carburetor>, consulted on December 18h, 2023.

back home and his wife asks him: "What was the car's problem?". John answers: "The problem was the carburetor". However, John does not know what a carburetor exactly is. Everything John associates with the term is "Some part of the car's engine". This description does not uniquely identify a specific part of the engine, as the engine includes other components. Therefore, according to an individualistic version of DT, John failed to specifically refer to his car's carburetor and his sentence is false or has no truth value. However, John arguably referred to his car's carburetor and truly said of it that it was the problem with his car. Even assuming that John's wife is aware that John is unable to provide a uniquely identifying description for "carburetor", she would be likely to interpret him as having successfully referred to the car's carburetor. For example, upon hearing her husband's utterance, she might begin calculating the money required for repairing the carburetor, contrasting it with the expenses necessary if the car issue had pertained to a different engine component. Hence, the description that determines the reference of John's use of "carburetor" is not the one that John possesses, but rather the one that experts entertain and that dictionaries document, reflecting the conventional meaning of the term.

One may object that also a strictly individualistic version of DT can account for John's successful reference to his car's carburetor and thus there is no need to posit any social version of the theory. John is aware that the information that he associates with "carburetor" is incomplete. Thus, John likely incorporates an additional specification into his description, to the effect that a carburetor is "a part of the car's engine with specific features *that experts are able to identify uniquely and precisely*". This sharpened description does pick out carburetors – albeit naturally by relying upon experts. Therefore, it does not seem necessary to move beyond an individualistic form of DT: it is indeed the description that John himself associates with the term that guarantees a successful reference to carburetors.[34] Crucially, this individualistic strategy does not seem to be viable in the Peano Case, where only a social version of DT can accommodate the prediction that the misinformed speaker's use of "Peano" refers to Peano. While in the Carburetor Case John does not believe that his information uniquely identifies any specific part of the engine, in the Peano Case the speaker, let us call him "Smith", genuinely believes that Peano is the discoverer of the axioms of arithmetic and considers

[34] This construal of DT alludes to some cursory considerations that Strawson develops (1959, p. 182n).

that information as uniquely picking out the name's referent. Therefore, in this circumstance, Smith's description arguably does not include any reference to experts: if someone asked Smith who Peano is, Smith would simply answer "The discoverer of the most common arithmetic axioms" and would take this description to successfully select one specific individual. Therefore, in the Peano Case, the only prediction that an individualistic version of DT can reasonably make is that John's use of "Peano" refers to Dedekind. DT can predict that John refers to Peano only under the assumption that what determines the referent of his use of the name is the description(s) that experts entertain.

If these considerations are correct, a case like the carburetor one that Reimer proposes fails to fully ground Machery's appeal to a social version of DT. To justify Machery's appeal to a social version of DT in the Peano Case, one needs an analogous case, not involving proper names, in which the speaker associates a term with a description that she herself takes to be complete and yet what determines the term's extension is not that description. Maybe a case along the following lines is suitable. Consider the term "bachelor": philosophers of language almost prototypically assume that a descriptive semantics covers this word. "Bachelor" means "unmarried man" and this description determines its extension. Most people know what the term means, and *a fortiori* socio-legal experts do. However, let us suppose that, due to a mishearing or misunderstanding, John associates the term with a wrong description, like "unmarried woman", and utters that "Bachelors are women". If John's description determined the extension of this occurrence of the term "bachelor", John's statement would be true. However, once again, the utterance is arguably false. Let us imagine that, during a university exam about socio-legal aspects of the law, John indeed claims that "Bachelors are women". The professor corrects him by stating that bachelors are in fact men and giving John's a lower grade. It seems a bit far-fetched to contend that John can make a case against his professor by emphasizing that, in light of the description that he himself associates with the term, what he said was true. Rather, if the student is intellectually honest, he would more likely acknowledge: "I *meant to claim* that unmarried women are female, but what I *actually* said is that unmarried men are female". The appeal to what he "actually" said as opposed to what he "meant" seems to adhere to the pragmatic/semantics distinction, namely, the divide between the literal and non-literal message of a sentence. These remarks suggest that what determines the literal message that John conveys is not the description that he

himself associates with the term "bachelor", but rather the description that socio-legal experts hold and that dictionaries report. It is possible to extend this social construal of DT to a descriptivist account of proper names, along the lines that Machery suggests. In the Peano Case, a speaker like Smith, who associates only "The discoverer of the arithmetic axioms" with "Peano", may still refer to Peano, and not Dedekind, by virtue of the descriptions that experts associate. In spite of Deutsch's claim to the contrary, the purported fact that "Peano" always refers to Peano does not necessarily constitute a counterexample against DT.

Naturally, there exists significant debate on the issue. For example, Devitt (2021, pp. 75-92) holds a different perspective on cases like John's. He marks a gap between literal and conventional meaning, and contends that, when John utters the statement "bachelors are women", he literally asserts that unmarried women are female. The problem with John's utterance is that its literal meaning does not align with the conventional meaning that the sentence "bachelors are women" conveys. John's answer in his test is wrong not because he asserted a literally false statement. Rather, the problem is that John was supposed to conform his idiolect with standard English, which he clearly did not, at least in the case of the term "bachelor".

In any case, much of the forthcoming discussion will not hinge on this dispute, as most of the experimental scenarios that I will discuss – both my own and the other authors' – specify that the speaker's description is the one that any other speaker – including experts – associates with the name. Therefore, any version of DT, individualistic or social, yields the same prediction regarding the name's referent, a prediction markedly different from the one that derives from CHT.

1.4.3 Linguistic usage

1.4.3.1 Martí's criticism

One may question Deutsch's argument by highlighting that Kripke appeals to intuitions also when discussing the Peano Case, and yet reinterpret Deutsch's stance as *normatively* suggesting that testing theories of reference necessitates a shift in focus from intuitions to referential facts. Devitt himself (2011, 2012b, 2015b), despite maintaining the higher value of experts' intuitions over laypeople's, claims that the most respectful approach to test theories on reference involves examining the reality these theories are about. Martí

(2009, 2012, 2014, 2020) identifies that reality in *linguistic usage*, that is, in how people use proper names. She argues that MMNS's study has semantical implications for theories of reference if it shows that East Asians tend to use names descriptively while Westerners causally-historically. However, according to Martí, MMNS's study would fail to provide any evidence in that regard. The question presented to participants asks about whom Johns talks when using the name "Gödel". A participant could engage with the question as follows. First, she considers the name in an abstract way. Second, she links it to its causal-historical chain on the one hand and to the speaker's description on the other. Finally, the participant expresses her intuition about the name's referent. A similar approach, however, does not provide any evidence about how the participant in fact uses names. Rather, it reflects how the participant *thinks* that speakers use names and reveals her theoretical inclination toward a specific view, be it causal-historical or descriptivist. In other words, the participant, based on a cursory examination of her overall experience with name usage, indicates whether she thinks that names form linkages with extra-linguistic entities via descriptions or causal-historical chains. Hence, it is not surprising that many participants express a descriptivist intuition. The idea that a name denotes the entity satisfying the associated description indeed has its credibility. After all, influential authors as Frege, Russell and Searle proposed and defended DT.[35]

1.4.3.2 Truth-value judgments

Machery, Olivola and de Blanc (MOD, 2009) reply to Martí's criticism by testing three groups, coming from India, France and Mongolia. MOD present a case similar to Gödel 2. As seen (section 1.3.2), in each group a part of the participants answer a RI question almost analogous to the one that Sytsma and Livengood (2011) use in their Clarified Narrator's Perspective. The other portion of participants answer this question:

> Having read the above story and accepting that it is true, when Ivy says, "Tsu Ch'ung Chih was a great astronomer", do you think that her claim is: (A) true or (B) false? (MOD, 2009, p. 690)

[35] The authors involved in the debate employ the terms "use" and "usage" interchangeably. I will adhere to this practice and I will take the two words as synonymous.

This is a truth-value judgment (TVJ) test, in which a participant has to evaluate whether a sentence is true or false. The *rationale* of the task is this: the speaker's inclination to judge a sentence as true indicates her inclination to pronounce that sentence, thus unveiling her linguistic usage. For example, if a person takes a sentence like "Snow is white" to be true, then she will be willing to utter "Snow is white".[36] Conversely, if she considers an assertion as false, for example "Snow is green", she will lack the inclination to utter it. The sentence "Tsu Ch'ung Chih was a great astronomer" is true only if "Tsu Ch'ung Chih" refers to the real discoverer of the scientific result. Hence, the answer "true" would show that the participant is willing to use the name "Tsu Ch'ung Chih" descriptively. Instead, the sentence is false if "Tsu Ch'ung Chih" refers to the other plausible referent, namely the thief. Hence, the answer "false" would show that the participant is willing to use the name causally-historically. The same reasoning applies, *mutatis mutandis*, to the name "Gödel" of Gödel 1.

Table 1.4 reports MOD's results in terms of causal-historical percentages of answers, with also the data regarding a group of American participants (drawn from Machery and Stich, 2012, but collected by Machery and Olivola).

Table 1.4 Percentage of causal-historical answers of MOD (2009) and Machery and Stich (2012)

Group	**TVJ**	**RI**
India	67	58
France	56	48
Mongolia	66	56
USA	64	67

The TVJ results corroborate the intra-cultural variation, as each group presents a non-negligible minority of participants providing the descriptivist answer. While the TVJ percentages are higher than the RI ones (except for the American sample), the difference is limited and never reaches statistical significance. In other words, the two values tend to be congruent. In light of this, MOD infer a cross-cultural variation across Americans

[36] A full elaboration of that principle would require some refinements, as MOD are aware: a TVJ relates to the participant's «desire to assert» a sentence, «though is not identical with it» (p. 691). For example, the principle does not extend to sentences including a personal pronoun. If Lionel Messi utters "I am Argentinian" and I take that sentence to be true, such a circumstance will not incline me to pronounce that sentence. In my utterance, "I" would refer to me, and I am not Argentinian.

and Chinese through an inductive reasoning. Given the above consistency of RIs and TVJs, the cross-cultural difference that MMNS (2004) detect would have emerged also if the Americans and Chinese participants had been asked to express TVJs, namely to perform a linguistic-usage test. As a consequence, MOD claim that Martí's objection does not neutralize MMNS's (2004) critical challenge.[37]

The inductive hypothesis that MOD formulate finds more direct corroboration with Li et al. (2018), who present two TVJ Gödel-style cases to American and Chinese participants, in English and Mandarin. Presented below is one of the cases, the Super Dog Race:

> Long ago, there was a famous dog race called the Super Dog Race. Max, Pickles, and Blaze were the dogs that joined the race among many others. Max and Pickles were two incredibly fast dogs. At the beginning of the race, they pulled out in front. All the way through the wilderness, they kept racing and racing. They left all the other dogs far behind. Max crossed the finish line first, winning the race. But Max got too excited. He couldn't stop running. He kept racing all the way to the North Pole. No one ever saw Max again. Pickles crossed the finish line second. He stopped and watched Max run away. The announcer of the race mistakenly thought that Pickles won the race. So he shouted, "Pickles, congratulations! You have won the Super Dog Race!". He then sent the news to every newspaper in the world that Pickles won the race. He also told them that another dog Blaze had run very fast even though he had short legs. Since then, people have all learned that Pickles won. But they don't know anything else about Pickles or the race. Tom and Emily are twins and they are in the first grade. They learned from their teacher at school that Pickles won the Super Dog Race. This is the only thing Tom and Emily know about the dog race and Pickles. They don't know anything about Max.
>
> The next day at breakfast, their dad asked: "Do you know who won the Super Dog Race?"
>
> Tom replied: "Blaze was the dog that won the Super Dog Race".
>
> Emily said: "Pickles was the dog that won the Super Dog Race".
>
> (Li, 2023a, p. 50)[38]

[37] An interesting video conversation between Genoveva Martí and Edouard Machery on the topic can be found at this link: <https://www.youtube.com/watch?v=a_DM292-JUs> (consulted on December 18th, 2023).
[38] The vignette is taken from Li (2023) because presented in its entirety. Li et al. (2018) solely report a «simplified version» of it (p. 109).

Participants express their TVJ by evaluating the statements that Tom and Emily make, determining whether they are «right» or not. Li et al. consider the judgment that Tom/Emily is right as equivalent to the TVJ "True", while the judgment that Tom/Emily is not right – namely, wrong – as equivalent to the TVJ "False". Li et al. do not explicitly provide reasons for preferring this approach instead of directly asking whether what Tom/Emily says is "true" or "false", but nothing substantial seems to hinge on that. Tom's assertion is clearly false, as it is a control statement. The critical statement is Emily's. According to DT, Emily's sentence is true, because what determines the reference of "Pickles" is the description "The dog that won the Super Dog Race". Hence, Emily is talking about Max, who won the race. According to CHT, the sentence is false because "Pickles" refers to the dog at the end of the chain of communication, and that dog did not win the race. Li et al. obtain 65% and 48% of causal-historical answers with Americans and Chinese, respectively. These findings confirm the inductive hypothesis that MOD formulate, indicating that Americans are more inclined to express a causal-historical TVJ than Chinese. Additionally, an intra-cultural variation within each group emerges as well.[39] A further TVJ study, Li (2021), corroborates the variation across American and Chinese participants. The experiment employs a new set of four vignettes and involves three conditions, differing in their final prompt. Of our interest at this stage is the final prompt in the "Original Condition" (I will discuss the other conditions in Chapter 3). It is

[39] The fact that Li et al. detect a cross-cultural variation is noteworthy in another aspect as well. Her Chinese sample comes from Mainland China, and therefore she presents the vignettes in simplified Chinese. This methodology marks a distinction between Li et al.'s experiment and the ones that I have discussed so far, which test participants from Hong Kong and use materials only in English (except for MSD's (2015) Clarified Gödel Case, which the author test also in traditional Chinese). The cross-cultural differences in RIs as well seem to persist regardless of the specific Chinese sample (whether from Mainland China or Hong Kong) and the language (English or simplified/traditional Chinese), as shown by Machery et al. (2010) – in reply to Lam (2010) – and Ding and Liu (2022). Moreover, although this dissertation does not delve into the developmental aspects of RIs and TVJs, a noteworthy finding in Li et al. (2018) concerns the emergence of cross-cultural variation also across American and Chinese children (aged seven): 60% of the American children choose the TVJ "False", as CHT predicts, while only 37% of the Chinese children do. More in general, the «cross-cultural results [on the reference of names across Americans and Chinese] are among the most robust findings of experimental philosophy» (Machery, 2015, p. 73; on the significance of demographic variations for experimental philosophy, see Machery, 2023; on the relevance of the replicability of experimental results, see Buckwalter, 2022; for a study evaluating the overall replicability of experimental philosophy, see Cova et al., 2018). Furthermore, van Dongen et al. (2020) conduct a meta-analysis of the experimental literature on proper names and conclude that the «meta-analysis supports the hypothesis that cross-cultural factors affect semantic intuitions about proper names» (p. 763). Van Dongen and colleagues also emphasize an «high inter-study variability of the data» (*ibidem*); however, as Machery (2024) highlights, that variability arguably emerges because van Dongen et al.'s work is «undiscerning with respect to the studies it meta-analyzes: Very different studies are aggregated, some of which have demonstrable flaws that have been discussed in the literature, and it is no wonder that van Dongen and colleagues find heterogeneous results. Some of their meta-analytic analyses aggregate results obtained with the Gödel and Jonah cases» (p. 190).

a TVJ prompt like the ones I have presented so far, except for the fact that – like in Li et al., 2018 – it elicits the TVJ by asking whether the speaker is right. Once again, Americans tend to choose the causal-historical option significantly more than Chinese. Specifically, American participants opt for the causal-historical TVJ 90% of the time, while Chinese participants 32%.

Li et al.'s (2018) Super Dog Race also serves to analyze a confounding factor that the Gödel-style scenarios may embed. One may observe a difference between the Gödel and Tsu Ch'ung Chih scenarios on the one hand and stories like for example the Super Dog Race on the other. This discrepancy lies in the moral quality of the actions that the characters perform. In the Gödel and the Tsu Ch'ung Chih stories, a man steals a scientific result from another person, constituting a blatant act of wrongdoing. Instead, in the Super Dog Race, the announcer mistakenly declares Pickles the winner, but Pickles himself does not commit any wrongdoing. Moreover, even assuming that Pickles witnessed Max crossing the line first and thus realized that he himself is not the real winner of the Race, Pickles – being a dog – is incapable of vocalizing an exculpatory statement like "There is a mistake, I'm not the winner of the race!". Although it is not *prima facie* clear why the moral valence of actions should affect how people understand proper names, one may wonder whether the moral variable somehow influences the results. The first scholars investigating that aspect are Beebe and Undercoffer (2015), who conduct RI studies in which they manipulate the moral value of some Gödel and Jonah-style vignettes. They conclude that the moral quality of the actions does not impact laypeople's RIs. Li (2023a, pp. 75-9) questions Beebe and Undercoffer interpretation of the results and carries out a TVJ study (Study 2) to further control for the moral valence of the vignettes. The results show that a vignette involving wrongdoing leads to an increase in the proportion of causal-historical TVJs among American participants, while a decrease among Chinese ones. Conversely, when the vignette does not involve any ethical violation, the difference in the TVJs between the two cultural samples diminishes. However, the effect is not robust and stable. Therefore, Li concludes that «moral evaluations only moderately influence reference resolution, thus playing little role in the formation of the cultural patterns of referential intuitions»[40] (pp. 150-1). The element of wrongdoing in the

[40] As Li's (2023) Study 2 is a TVJ one, by «referential intuitions» Li means "TVJs", and not the referential judgments like the ones at the center of MMNS (2004).

classical Gödel and Tsu Ch'ung Chih stories does not exert a substantial enough impact to be a worrying confounding factor.

1.4.3.3 TVJs are epistemically ambiguous

Vignolo and Domaneschi (2018) and Domaneschi and Vignolo (2020) do not agree on the reasons that MOD provide for the claim that TVJs capture linguistic usage. The circumstance that a participant judges the sentence "Gödel is a mathematical genius" uttered by John as true does not entail that she will be inclined to utter that sentence. The experimental subject reads the entire vignette and is aware of the real course of the events. Since the vignette presents Gödel as the thief who stole the theorem, the subject will associate "Gödel" with a description like "The man who stole the theorem and claimed credit for it". In this context, DT as well and not only CHT predicts that the participant, having accepted the vignette's story as true, will use the name "Gödel" to talk about the thief. In a word, John finds himself in a state of "semantic contradiction"[41] between the description that he associates with the name and the individual causally-historically connected to it. Instead, the participant is not in such a state.[42]

Nonetheless, Domaneschi and Vignolo assert that there are other reasons for claiming that TVJs capture linguistic usage. Domaneschi and Vignolo follow Devitt and Porot (2018) and contend that TVJs shed light on the participant's *understanding* of the sentence. One's linguistic competence in using names unfolds not only in the production of sentences containing them, but also in the comprehension of names when appearing in the sentences that one's interlocutor utters. To echo Cohnitz's words: «presumably, production and interpretation are just two sides of the same coin» (2015, p. 96; see also Cohnitz and Haukioja, 2013, 2015; Maynes, 2015). If a participant judges the sentence "Gödel is a mathematical genius" uttered by John as true, that reveals that the participant understood the sentence to express a proposition about Schmidt. That is evidence about

[41] I take this elegant expression from Domaneschi, Vignolo and Di Paola (2017, p. 3).

[42] This criticism aligns with the one that Devitt and Porot (2018) term the "New Meaning Objection" and that I will discuss in Chapter 2 extendedly. To be precise, the New Meaning Objection is not the main reason Vignolo and Domaneschi (2018) and Domaneschi and Vignolo (2020) reject of the thesis that TVJs unveil linguistic production. Vignolo and Domaneschi's (2018) principal argument centers around the potential semantic/speaker reference ambiguity in the Gödel Case (a thesis that they later reject in their Vignolo and Domaneschi, 2022). Domaneschi and Vignolo (2020) assert that TVJs cannot serve as complete tests of usage due to their inherent epistemic ambiguity, a thesis that I extensively examine in this section. Although Domaneschi and Vignolo do not explicitly emphasize it as their primary criticism, they do endorse the New Meaning Objection, as evident from Vignolo and Domaneschi (2018, pp. 8-9, n. 7) and Domaneschi and Vignolo (2020, p. 445, n. 8).

linguistic usage and not RIs, because the participant is not expressing any judgment like "When using 'Gödel', John is talking about Schmidt", nor is there evidence that she is grounding her TVJ upon it. In light of this, the TVJ "True" supports DT. For opposite reasons, the TVJ "False" corroborates CHT.

At this point, Domaneschi and Vignolo may seem to take MOD's study to neutralize Martí's objection, albeit for different reasons than those that MOD advocate. However, Domaneschi and Vignolo argue that TVJ tests can capture linguistic usage effectively only if they are free from the epistemic ambiguity, which poses a threat not only to RIs, as Sytsma and Livengood (2011) note, but also to TVJs. Specifically, the epistemic ambiguity pertains to the interpretation of the truth predicate. According to Domaneschi and Vignolo, the predicate "true" can mean "true according John's perspective", i.e. "true according to the perspective of the vignette's character", or "True according to the narrator's perspective", i.e. "factually true". Once again, only the narrator's perspective elicits the collection of TVJs that are adequate to compare the two classical theories of reference.[43]

A participant may interpret "Gödel" as referring to Gödel – the thief – and nevertheless answer that John's statement "Gödel is a mathematical genius" is true. That occurs if the participant understands "true" as "true from the standpoint of the vignette's character". The sentence is indeed true from John's epistemic perspective, because it reflects the consensus within the linguistic community to which he belongs.[44] Regardless of how a participant understands "Gödel", descriptively or casually-historically, it is beyond question that John *believes* that what he is saying is true. After all, what teachers told John at school is that "Gödel is the author of an important scientific result, the incompleteness theorem". Therefore, if the participant understands the predicate "true"

[43] Similarly to the discussion regarding RIs, in this case as well I harbor some reservations regarding the use of the labels "ambiguity" and "narrator's perspective". I will explore this topic in section 2.2.2.2.

[44] One may object that it is somewhat imprecise to claim that a participant adopting the speaker's perspective will judge John's sentence as true *because* John's claim aligns with the consensus within his community. One matter is the perspective, i.e. the belief, of the single speaker (John), while another is the perspective, i.e. the consensus, emerging within a broader group of people (like a community). The former does not always coincide with the latter: an individual may have a belief different from the one that the majority of her community holds. For example, John may believe that apples are not healthy, while the majority of the people of his community may believe that they are: in such a circumstance, John's belief diverges from the consensus prevailing in his community. I acknowledge and share these concerns, and I will explore them in section 2.3.2.2. That said, for the moment, I will consider the labels "speaker's perspective" and "community's perspective/consensus" as interchangeable: this choice is inconsequential, since in all the scenarios under discussion what the speaker believes aligns with the consensus within her own community. I thank Genoveva Martí for precious reflections on this topic.

from John's point of view, the TVJ sheds no light on the participant's linguistic usage and does not contribute to testing theories of reference.

The TVJ is relevant only if the participant understands the truth predicate as "true strictly speaking". The subject needs to assume the viewpoint of the omniscient narrator, who possesses some privileged knowledge inaccessible to John. Specifically, the narrator is aware that a theft happened and thus the description that John associates with the name does not pick out the original bearer of the name. Adopted that perspective, the participant judges whether what John is asserting, on top of being believed as true by John, is also *factually true*. That is the point of disagreement between DT and CHT: the sentence is true according to DT, but not CHT. MOD do not consider the epistemic ambiguity. Therefore, Domaneschi and Vignolo argue that MOD expose themselves to a type 2 error when inferring the congruency between use and intuitions.

Domaneschi and Vignolo (2020) conduct three experiments to disambiguate the epistemic perspective. At this stage, I will focus on their Experiment 1 (in Chapter 2 I will discuss Experiment 2 and 3). Domaneschi and Vignolo test a group of Italian participants with an Italian translation of the simplified version of the Super Dog Race that Li et al. reproduce in their (2018). The text is the following:

> Long ago, there was a race called the Super Dog Race. Max, Pickles and Blaze participated in the race. Max crossed the finish line first, winning the race, but he got too excited and ran all the way to the North Pole. Pickles crossed the finish line second. He stopped and watched Max run away. The announcer of the race mistakenly thought that Pickles won the race. He told every newspaper in the world that Pickles won. He also told them that another dog, Blaze, ran very fast despite his short legs. Since then, everyone learned that Pickles won the race. They don't know anything else about Pickles.
>
> Tom and Emily learned at school that Pickles won the Super Dog Race. This is the only thing they know about the dog race and Pickles. They don't know anything about Max.
>
> (Domaneschi and Vignolo, 2020, pp. 448-9)

Participants are asked to perform a TVJ judgment:

> That night, their dad asked: Do you know who won the Super Dog Race?
> Emily said: "Pickles was the dog that won the Super Dog Race".
> Do you think that her claim is
> (a) true or
> (b) false?
> (Domaneschi and Vignolo, 2020, p. 449)

Those participants who choose "True" – the purportedly descriptivist answer – are asked a follow-up question. The aim is to ascertain the epistemic perspective that they used when providing the TVJ.

> You think Emily's claim is true because:
> (a) Pickles did not win the Super Dog Race, but Emily believes that Pickles won the race because her teacher told it to her.
> (b) What Emily believes is a true description of Max and she uses the name "Pickles" to talk about the dog that really (unbeknown to Tom and Emily's teacher) won the Super Dog Race.
> (Domaneschi and Vignolo, 2020, p. 449)

If a participant judges the sentence that Emily utters is true because she understands the name descriptively, then she should choose (b), for (a) pertains to what John believes, while (b) to what John asserts beyond what he believes. Domaneschi and Vignolo observe that 42% of participants answer that the sentence is true. However, 75% of them then answer (a). The majority of the participants provide a TVJ that is only seemingly descriptivist, as the follow-up question shows that they adopt the character's epistemic viewpoint. Therefore, the epistemic ambiguity impacts the group of Italian participants whom Domaneschi and Vignolo test. For the most part, the supposedly descriptivist answers actually provide no evidence in support of DT. This ambiguity can have a bearing also on the participants of the various nationalities of MOD's experiment, and thus the data that MOD present fail to justify the thesis of an intra and/or cross-cultural variation in the use of names.

The aforementioned Original Condition in Li (2021) yields similar results. Li presents a set of four vignettes to American and Chinese participants (in English and Mandarin). The former provide the TVJ that CHT predicts 90% of the time, while the latter 32% of the time. Subsequent to each TVJ, participants answer an open-ended question asking them the rationale behind their TVJ. Li finds that, when Americans answer "True" (a strict minority of the time) and Chinese do that (a large majority of the time), thus providing the TVJ that DT predicts, they ground their explanations on the character's epistemic perspective. Their justificatory remarks revolve around what Emily – the speaker – has been told, what is usually said within her linguistic community and thus what she believes to be true. Notably, Li does not collect any genuinely descriptivist explanation, i.e., answers explicitly stressing that Emily is right because she talks about the entity that meets the information that she associates with the name. Therefore, the epistemic ambiguity also impacts on American and Chinese participants. Concerning the TVJ "False", both Americans and Chinese explain their answer by appealing to the narrator's viewpoint, that is, the perspective in which scholars are interested since it elicits conflicting predictions by DT and CHT. In light of this, the different percentages of TVJs "True" and "False" across Americans and Chinese participants do not seem to arise from different patterns in the use of names. Rather, the variation seems to stem from a different inclination across the two populations to adopt either the narrator's viewpoint or the character's. Li claims that independent evidence corroborates such an explanation, as studies like Wu and Keysar (2007), Luk et al. (2012) and Wu et al. (2013) show how Chinese adopt other people's perspective more efficiently than Americans.[45]

[45] Cohnitz (2015, p. 106, n. 16) is the first author suggesting that Wu and Keysar's (2007) results might potentially constitute the starting point for explaining the cross-cultural variation that MMNS (2004) unveil. Wu and Keysar conduct an eye-tracking study in which the participant is in front of a confederate, who acts as an instructor. Some objects are visually accessible to both the participant and the instructor, while other objects are visually accessible only to the participant and the confederate is not even aware of their existence (as far as the participant knows, of course). At this point, the confederate asks the participant to relocate certain objects. Clearly, the confederate can only intend to refer to the objects that are within the mutual visual field. However, some of the phrases that the confederate uses crucially happen to have the hidden objects as their best or at least plausible referents. When the instructor uses those expressions, Americans turn out to fixate their gaze upon the interfering hidden object statistically significantly more than Chinese, before directing their gaze toward the object that is best satisfier among those within the common visual field. This finding suggests that, compared to Americans, Chinese can take the other's perspective more easily, thus demonstrating a willingness to prioritize the shared knowledge over their own privileged awareness of concealed objects. This central finding receives further experimental corroboration from Luk et al. (2012) and Wu et al. (2013).

1.4.3.4 Elicited production

Devitt and Porot (2018) present another test of linguistic usage, which focuses on linguistic production, namely the sentences that speakers utter. If anything is linguistic usage, then production for sure is. For example, suppose that an experimenter is interested in the linguistic usage of the word "table": the most natural way to investigate it seems exactly to ask participants to employ the term and observe the sentences they utter. While criticizing RI tests, Martí (2009) suggests that the most effective method to study the linguistic usage of names is to ask participants to produce sentences containing them. For example, Martí suggests a prompt aimed to elicit participants' linguistic production:

> One day, the fraud is exposed, and John exclaims: «Today is a sad day: we have found out that Gödel was a thief and a liar».
> What do you think of John's reaction?
> (Martí, 2009, p. 47)

According to DT, "Gödel" refers to the real author of the theorem, Schmidt. Therefore, if a participant is a descriptivist user of proper names, she should react by stressing that John's assertion is irrational. She should articulate something akin to the following: "What John says does not make sense! Gödel truly discovered the theorem". Therefore, «after so many years of being robbed of his due credit, John is now adding insult to injury by declaring the poor man a liar and a thief!» (p. 47). Instead, John's assertion should appear rational to a causal-historical user of names, as John is talking about the original bearer of the name, who – in the fictional story – is a thief and a liar.

Devitt and Porot present a case modeled after Gödel 2 to a group of American participants:

> Students in astronomy classes in Hong Kong are told that a man called "Tsu Ch'ung Chih" first determined the precise time of the summer and winter solstices. This is the only thing that typical Hong Kongers ever hear about this man. Now suppose that that man did not make the discovery he is credited with. He stole it from an astronomer who died soon after making the discovery. But the theft remained

entirely undetected and so the man that Hong Kongers have been told about became famous for the discovery of the precise times of the solstices.
(Devitt and Porot, 2018, p. 1562)

In a between-subjects design, participants are presented with different final prompts, soliciting the participants' RIs, TVJs or elicited production (EP). For the reasons above, Devitt and Porot take EP to be «a *pure* test of usage» (p. 1559, my emphasis). Devitt and Porot consider TVJs as well as tests of linguistic usage, albeit «somewhat imperfect» (*ibidem*). A TVJ «"puts words into the mouth" of the participant» (p. 1561), since the subject evaluates the truth value of the sentence that the experimenter provides rather than freely producing her own. Given that EP tests are the gold benchmark for linguistic usage, the extent to which RI and TVJ results align with the causal-historical or descriptivist patterns that the EP data provide determines the degree to which those tests effectively reflect linguistic usage. The final prompts are the following:

EP
Having read the above story and accepting that it is true, what is your opinion of Tsu Ch'ung Chih?

TVJ-DT
Having read the above story and accepting that it is true, please indicate below whether you think the following statement is true or false.
"Tsu Ch'ung Chih was a great astronomer".

TVJ-CHT
Having read the above story and accepting that it is true, please indicate below whether you think the following statement is true or false.
"Tsu Chu'ung Chih was a thief and a liar".

RI

Having read the above story and accepting that it is true, who does the name "Tsu Ch'ung Chih" refer to?

(a) The man who discovered the summer and winter solstices.

(b) The man who stole the discovery and took credit for it.

(Devitt and Porot, 2018, pp. 1565-66)

TVJ-DT is a TVJ that is true according to DT and false according to CHT. Predictions are the other way around with TVJ-CHT. Devitt and Porot use two TVJs to control for the participants' acquiescence bias to respond "True" (see Krosnick, 1999, pp. 552-3).[46]

In the EP test, participants express a positive or negative opinion about Tsu Ch'ung Chih. Let us assume that the participant formulates a negative judgment, for example "Tsu Ch'ung Chih is a thief and impostor". Devitt and Porot interpret that answer to provide corroboration to CHT, according to which the referent of "Tsu Ch'ung Chih" is the thief. On the other hand, if the judgment is positive, such as "Tsu Ch'ung Chih is a great scientist", then Devitt and Porot construe the participant as providing an answer in

[46] The acquiescence bias is such as to lead to an «inflation of rates of agreement with claims like the target prompt, irrespective of their content» (Adams and Hansen, 2020, p. 111). That is, especially in the context of a statement with which participants neither strongly agree nor disagree, they typically lean more toward agreement. In the context of Devitt and Porot's experiment, this bias would reflect in a higher number of TVJs "True" compared to the amount of TVJs "False". Therefore, if for example participants turn out to overall agree with a sentence P that CHT predicts as true, one may object that the result may be due to the participants' bias toward answering "True" rather than to a genuine causal-historical commitment. However, if that result lines up with causal-historical results concerning a statement Q that DT predicts as true, then that combined finding provides corroboration to CHT: what guides the participants' agreement with P is arguably a causal-historical commitment rather than an acquiescent tendency.

line with DT, according to which the referent of "Tsu Ch'ung Chih" is the real discoverer of the astronomic result. Table 1.5 reports the results.

Table 1.5 Results of Devitt and Porot (2018)

Prompt	**Percentage of causal-historical answers**
EP	95
TVJ-DT	95
TVJ-CHT	81
RI	80

Devitt and Porot observe a clear tendency to provide causal-historical answers, which is particularly noteworthy in the case of the EP test due to its innovative nature within the literature. Almost all participants, 95%, express a negative judgment about Tsu Ch'ung Chih, and thus these results seem to constitute compelling evidence in support of CHT.

Devitt and Porot anticipate a possible criticism against their study. Since the participants read the vignette containing the complete and privileged account of the theft, one may argue that they will associate the name "Tsu Ch'ung Chih" with "The person who appropriated the proof of the discovery of the solstice times through an entirely undetected theft". That description denotes the same person whom the causal-historical chain identifies as the name's referent. Hence, both DT and CHT predict that the participant will use the name to talk about the thief. Devitt and Porot call that criticism "The New Meaning Objection", because for the descriptivist participant the name "Tsu Ch'ung Chih" would acquire a new meaning alongside the already established one within the Hong Kong community. Devitt and Porot reply to the objection experimentally, by conducting a follow-up EP study.[47] The final prompt now is: "Having read the above story and accepting that it is true, what would you say to a Hong Konger about Tsu Ch'ung Chih? (Please write as if you are speaking to a Hong Konger)" (p. 1573). This

[47] As seen, the acceptance of the New Meaning Objection is the reason Vignolo and Domaneschi (2018) and Domaneschi and Vignolo (2020) deny that the TVJs in MOD (2009) unveil the sentences that participants would utter. Devitt and Porot also advance some theoretical reasons for the claim that The New Meaning Objection is implausible: I will delve into them in section 2.

revised final prompt specifies that the interlocutor is a Hong Konger, and the parenthetical remark "Please write as if you are speaking to a Hong Konger" further stresses that the participant needs to formulate her answer with consideration for the recipient. The prompt is intended to guide the descriptivist participant to conform her use of the name to the description that the members of the Hong Kong community associate with it. The results of this follow-up experiment still provide unambiguous support to CHT: 89% of participants formulate a judgment in line with the Kripkean theory. The value of the follow-up results depends on the authentic clarificatory efficacy of Devitt and Porot's instruction to write as if one were speaking to a Hong Konger. Machery (2021) raises doubts as to whether that specification compels or significantly inclines the descriptivist participants to use "Tsu Ch'ung Chih" to talk about the real discoverer of the scientific result. I will thoroughly examine the New Meaning Objection and the value of Devitt and Porot's follow-up experiment in section 2.[48]

Devitt and Porot specifically address the New Meaning Objection only in relation to their EP test, likely because linguistic production constitutes the central and innovative methodology of their study. However, the objection arguably extends to their RI and TVJ tests as well. A key distinction between Devitt and Porot's RI and TVJ tests and those of previous literature lies in the absence of the speaker, for example Ivy or John. As a way of example, let us consider MOD's (2009) formulation of the RI and TVJ final prompt. Their RI final prompt reads as follows: «When Ivy uses the name "Tsu Ch'ung Chih", who do you think she is actually talking about?» (MOD, 2009, p. 690). Ivy is the utterer also in their TVJ tests: «When Ivy says, "Tsu Ch'ung Chih was a great astronomer", do you think that her claim is: (A) true or (B) false?» (*ibidem*). Ivy belongs to the linguistic community with a limited epistemic perspective on the course of the events. The only description that Ivy associates with "Tsu Ch'ung Chih" is "The discoverer of the solstice times". Hence, as per DT, that is the only meaning according to which Ivy can use the

[48] Devitt and Porot (2018) is neither the only nor the first study testing the reference of proper names through linguistic production. Domaneschi, Vignolo and Di Paola (2017) present some Gödel cases (in Italian) and then ask participants to choose among various options to complete some sentences about the scenarios. The option that the participant select reflects the words that she would utter to complete the sentence. Hence, Domaneschi et al.'s study is a semi-elicited production test, where the word "semi" signals that the participant does not freely write her own words, but rather chooses from a predetermined set of possible options. The authors collect causal-historical results. Nevertheless, that study too is susceptible to the New Meaning Objection: the participant will complete the sentence based on the information that she herself associates with the name. I have presented Devitt and Porot (2018) for two reasons. First, it is a test of elicited production, rather than semi-production. Second, the authors acknowledge the New Meaning Objection and address it through a follow-up study.

name “Tsu Ch’ung Chih”. However, in Devitt and Porot’s tests, Ivy is not present. Therefore, contingent upon the acceptance of the New Meaning Objection, the descriptivist participants can express their RIs or TVJs according to the new meaning that they possess.[49]

1.5 The literature on the Jonah Case

The literature on the Jonah Case is relatively less extensive than the one on the Gödel Case. The discrepancy is likely due to the findings of MMNS (2004), who identify an intra-cultural variation, but not a cross-cultural one, in the context of the Jonah Case. In this section, I will review the experimental literature on the Jonah Case.

1.5.1 RI studies

Let us rehearse the vignettes that MMNS (2004) employ in their investigation of the Jonah Case. I will report the one using Western names, but the forthcoming considerations also apply to the vignette featuring East Asian names:

> In high-school, German students learn that Attila founded Germany in the second century A.D. They are taught that Attila was the king of a nomadic tribe that migrated from the east to settle in what would become Germany. Germans also believe that Attila was a merciless warrior and leader who expelled the Romans from Germany, and that after his victory against the Romans, Attila organized a large and prosperous kingdom.
>
> Now suppose that none of this is true. No merciless warrior expelled the Romans from Germany, and Germany was not founded by a single individual. Actually, the facts are the following. In the fourth century A.D., a nobleman of low rank, called “Raditra”, ruled a small and peaceful area in what today is Poland, several hundred miles from Germany. Raditra was a wise and gentle man who managed to preserve the peace in the small land he was ruling. For this reason, he quickly became the main character of many stories and legends. These stories were passed on from one generation of peasants to the next. But often when the story was passed on the

[49] The first author suggesting the elimination of the speaker (John/Ivy) for TVJ tests is Wikforss (2017, p. 106). According to Wikforss, this exclusion brings researchers closer to linguistic-production data, as the TVJ that the participant provides no longer reflects whether she takes some *other* speaker to express a true proposition, and thus the TVJ more directly unveils whether the participant herself would be disposed to utter the sentence. However, Wikforss does not consider the New Meaning Objection.

peasants would embellish it, adding imaginary details and dropping some true facts to make the story more exciting. From a peaceful nobleman of low rank, Raditra was gradually transformed into a warrior fighting for his land. When the legend reached Germany, it told of a merciless warrior who was victorious against the Romans. By the eighth century A.D., the story told of an Eastern king who expelled the Romans and founded Germany. By that time, not a single true fact remained in the story.

Meanwhile, as the story was told and retold, the name "Raditra" was slowly altered: it was successively replaced by "Aditra", then by "Arritrak" in the sixth century, by "Arrita" and "Arrila" in the seventh and finally by "Attila". The story about the glorious life of Attila was written down in the eighth century by a scrupulous Catholic monk, from whom all our beliefs are derived. Of course, Germans know nothing about these real events. They believe a story about a merciless Eastern king who expelled the Romans and founded Germany.

When a contemporary German high-school student says "Attila was the king who drove the Romans from Germany", is he actually talking about the wise and gentle nobleman, Raditra, who is the original source of the Attila legend, or is he talking about a fictional person, someone who does not really exist?

(a) He is talking about Raditra.

(b) He is talking about a fictional person who does not really exist.

(MMNS, 2004, B9-B10)

As seen in section 1.2, the answers of each participants can range between 0 and 2. "0" corresponds to a descriptivist answer to both versions of the Jonah Case, "1" to a causal-historical and a descriptivist answer and "2" to a causal-historical answer to both vignettes. Westerners' mean score is 1.23 (SD: 0.96), while East Asians' is 1.32 (SD: 0.76). The difference between the two groups is not statically significant. Both groups tend toward the answer predicted by CHT.

Some flaws in the wording of the vignette arguably compromise the overall reliability of the results. First, MMNS employ two versions of the same name: "Raditra" and "Attila". The former is the original name of the nobleman of low rank. The latter is the name that contemporary people inherit due to variations of the name over time. According to CHT, the name does not need to retain a single and constant orthographic form throughout history to be considered the same name. For example, the way we currently write and pronounce "Jonah" differs from ancient times. Yet, not for this reason did the

communication chain get broken; indeed, “Jonah” is one of the main examples that Kripke himself uses. In other words, from a Kripkean perspective, the crucial element is the chain of communication, not the retention of a single orthographic from. Therefore, in Jonah 1, CHT entails that “Attila” refers to the individual who was the bearer of the name from which “Attila” derives; that is, CHT does entail that “Attila” refers to Raditra. Hence, my criticism is not on a theoretical level. Rather, my criticism is on an experimental level: the use of two versions of the same name arguably introduces a further and probably unnecessary level of complexity, a factor that researchers should avoid when testing laypeople’s intuitions. As the alteration of the name over time is not essential to the vignette, it is possible to simplify the story by dropping it.

Second, the vignette length presents a notable challenge, as MMNS (2004) acknowledge: «setting out the Jonah cases precisely requires a lengthy presentation […], so it is possible that our probes were simply too long and complex to generate interpretable data» (p. B7). The formulation of a Gödel scenario requires the misattribution of some property, such as being the author of the incompleteness theorem. Instead, the formulation of a Jonah scenario requires the researcher to come up with a description that the original bearer of the name and *any* other individual *unequivocally* fail to satisfy. According to DT, the referent of a name is the satisfier or the *best* satisfier of the description(s): thus, an individual can be the name’s referent even if she fails to satisfy a part of the properties associated with the name. Hence, a researcher needs to conceive a description whose central aspects are *entirely or almost entirely false* of any individual. To accomplish this task, a higher number of words may be necessary compared to those that one needs to merely misattribute a single property.

Third, let us consider the descriptivist option: «He [John] is talking about a fictional person who does not really exist» (p. B10). As Jackman (2009, p. 169) notes, «what [MMNS] treat as the ‘descriptivist’ doesn’t posit reference failure as much as it does the speaker referring to a fictional character». For example, a descriptivist option stressing the reference failure could be: “John is talking about no person at all”. The predicament with referring to a fictional character lies in the challenge that laypeople may encounter in comprehending entities, such as fictional ones, that fundamentally lack real existence. To echo Beebe and Undercoffer’s (2016) words: «the very idea of talking about things

[i.e., fictional entities] that aren't things at all (because they do not exist) might cause some processing difficulties with philosophically untrained participants» (p. 350).

Beebe and Undercoffer (2015) is the first study investigating the Jonah case after MMNS (2004). As already mentioned in 1.4.3.2, Beebe and Undercoffer intend to explore some previously uncontrolled for moral factors of the Gödel and Jonah cases as MMNS (2004) present them. The Gödel Case involves an element of wrongdoing, as Gödel claimed the credit for the result by dishonestly appropriating it from Schmidt. Instead, the Jonah case lacks any such element. The central character of the vignette – Attila/Raditra – is even a morally commendable individual, as people started to develop legendary tales of him because «Raditra was a wise and gentle man who managed to preserve the peace in the small land he was ruling» (MMNS, 2004, p. B10). Research in the field of experimental philosophy highlighted that moral considerations can influence laypeople's intuitions regarding various matters. For example, Knobe (2003) suggests that the ascription of intentionality to side effects depends on the moral quality of the effect itself. Beebe and Undercoffer speculate whether the influence of moral valence might similarly affect participants' RIs as well, potentially explaining the observed cross-cultural variation and the different pattern of results across the Gödel and Jonah cases. Put differently, the divergence in RIs between American and Chinese participants in response to the Gödel Case may be due to their distinct attitudes toward the misconduct within it, while the lack of cross-cultural variation in the Jonah case would be due to the absence of any wrongdoing. To use Beebe and Undercoffer's words: «if there are cross-cultural differences in sensitivity to moral violations, we reasoned that this might have a distorting effect on participants' intuitions about these cases» (p. 450).

In light of these considerations, Beebe and Undercoffer present two Jonah cases, one with Western names and the other with East Asian names, to participants from the US and China. The two Jonah cases involve elements of wrongdoing. A couple of additional features merit attention. First, in contrast to the Jonah cases that MMNS (2004) outline, Beebe and Undercoffer do not use a name undergoing a process of alteration over time. Second, the descriptivist option does not posit reference to a fictional entity. The option states that the speaker «is not talking about Raditra's peace-loving grandfather» (p. 452), who in the vignette's story is the original bearer of the name. As the vignette does not present any individual as a good candidate to be the best satisfier of the tales associated

with the name, the descriptivist option in fact outlines a genuine reference failure. Under these two aspects, Beebe and Undercoffer's Jonah cases represent an improvement of the ones that MMNS (2004) use.

Westerners and East Asians express the causal-historical RI 39% and 34% of the time, respectively. The difference is not statistically significant. In comparison to MMNS (2004), the results confirm that, while a cross-cultural variation does not emerge, an intra-cultural does. In light of this, Beebe and Undercoffer conclude that «moral valence has not had a distorting factor on participants' responses to the kinds of cases that have been used thus far in the literature» (p. 450). The conclusion is surely justified insofar as both MMNS and Beebe and Undercoffer fail to detect a cross-cultural variation and identify only an intra-cultural one. However, in MMNS (2004) both groups – US and China – lean toward CHT, while in Beebe and Undercoffer (2015) toward DT. It is difficult to determine whether the difference is due to the moral variable altering across the two studies. As mentioned, Beebe and Undercoffer's vignettes differ from MMNS's not only in terms of ethical aspects, but also for the name's retention of a single orthographic form and for the inclusion of a descriptivist option that posits genuine reference failure.

In a subsequent work, Beebe and Undercoffer (2016) expand upon the understanding of the Jonah Case through two studies. The first, Study 1a, involves testing a group of American and Chinese participants by employing the same Jonah vignettes and final prompts of MMNS's (2004) research. Americans and Chinese express the causal-historical RI 39% and 25% of the time, respectively. The difference is statistically significant and therefore the expected cultural divergence receives corroboration: although both groups tend toward the descriptivist answer, this inclination is stronger within the Chinese sample. Another Jonah experiment in Beebe and Undercoffer (2016), Study 2b, further corroborates the overall trend. The authors present two Jonah Cases analogous to those in MMNS (2004), with the exception that the descriptivist option does not posit reference to a fictional entity. Like in Beebe and Undercoffer (2015), the descriptivist alternative simply states that the name does not refer to the original bearer, as it reads – in its Western-name version – «He is not talking about Raditra» (p. 351). The results reveal that Americans and Chinese express the causal-historical RI 47% and 37% of the time, respectively. The difference is again statistically significant.

Beebe and Undercoffer's (2015, 2016) studies provide interesting insights on the RIs on the Jonah Case. As regards the general support for DT and CHT, both MMNS's (2004) groups – American and Chinese – tend toward CHT, despite a non-negligible intra-cultural variation. However, the three Jonah studies that Beebe and Undercoffer (2015, 2016) conduct present a different pattern, showing a general alignment with the descriptivist predictions. As regards the cross-cultural variation, both MMNS (2004) and Beebe and Undercoffer (2015) fail to find any evidence of it in the Jonah Case. Instead, the two studies in Beebe and Undercoffer (2016) reveal a cross-cultural variation, aligning with the expected trend of Americans leaning more toward CHT than Chinese. These results suggest some level of cross-cultural variation in the Jonah Case, although it may not be as robust as the one emerging in the context of the Gödel Case.

Islam and Baggio investigate the RIs on the Jonah Case with two experiments. As explained (section 1.3.1, n. 19), they test both the Gödel and the Jonah Case. However, their reading of the outcomes related to the Jonah cases is debatable; as a consequence, the interpretation of their statistical analyses is not straightforward, as they are based both on the Gödel outcomes and the Jonah ones. Islam and Baggio contend that, in a scenario such as Jonah 1, CHT predicts that "Attila" has no referent, because no man is the original bearer of that name – the individual at the end of the chain was known as "Raditra".[50] While I share some concerns regarding the use of two orthographic forms, my criticism – as previously explained – is on an experimental level, not on a theoretical one: CHT does imply that "Attila" refers to Raditra. Therefore, I will refrain from reiterating my arguments to the reader. However, from their raw data[51], one can derive the specific Jonah Case percentages. In their Experiment 1, Islam and Baggio administer three Jonah-style vignettes, two of them being overall similar to MMNS's Jonah 1 and Jonah 2. The main differences lie in the length of the text (Islam and Baggio shorten the vignette) and in the

[50] Therefore, according to Islam and Baggio, the predictions of DT and CHT concerning the Jonah Case coincide, as both theories would predict that the name fails to refer; instead, a speaker-reference interpretation of the name would predict that "Attila" refers to Raditra. I will come back to the overall design and rationale of their experiment in section 3.1.1.

[51] Accessible at <https://doi.org/10.1093/jos/ffaa001>.

structure of the final prompt. The final prompt of Islam and Baggio's vignette resembling Jonah 1 is as follows[52]:

> When a contemporary German high school student uses the name 'Attila', she is talking about:
> (A) The German warrior who expelled the Romans from Germany.
> (B) A fictional person who does not really exist.

Instead, in the original MMNS (2004), the final prompt of Jonah 1 reads:

> When a contemporary German high-school student says "Attila was the king who drove the Romans from Germany", is he actually talking about the wise and gentle nobleman, Raditra, who is the original source of the Attila legend, or is he talking about a fictional person, someone who does not really exist?
> (A) He is talking about Raditra.
> (B) He is talking about a fictional person who does not really exist.
> (MMNS, 2004, p. B10)

Islam and Baggio modify the final prompt for the Jonah Case, bringing it into closer alignment with the wording of the Gödel cases. Islam and Baggio eliminate any explicit sentence uttered by the contemporary German high school student: they simply ask about whom she talks "when using the name 'Attila'". Moreover, the causal-historical option selects the Kripkean referent through a description rather than a name (e.g., "Raditra"): that is in line with the approach of the Gödel scenarios, where the prompt presents the thief as "the person who got hold of the manuscript and claimed credit for the work" rather than as "Schmidt". Those modifications arguably represent an improvement in design, ensuring greater consistency across the formulations of the Gödel scenarios and the Jonah ones.

Across the sample of three Gödel-style vignettes, administered to each participant, the percentages of causal-historical answers ("The German warrior who expelled the Romans from Germany") are as follows: 60% for the Norwegian group, 77% for the immigrant group and 85% for the participants born and living in Bangladesh. The causal-historical

[52] Drawn from the supplementary material, accessible at <https://doi.org/10.1093/jos/ffaa001>.

percentages for the Norwegian group are numerically lower than those for the Asian groups.

In Experiment 2, Islam and Baggio run a similar experiment, with both the Gödel and the Jonah Case, with Norwegians and people living in Bangladesh. Islam and Baggio administer to each participant two out of a sample of three Jonah scenarios. In this case, though, the materials are not in English, but in Bokmål Norwegian and Bangla. Aggregating the RIs related to the Jonah cases, Norwegians express the causal-historical RI 70% of the time while Bangladeshis 93%. Therefore, also in Experiment 2 the Norwegian group exhibits a lower inclination for the Kripkean answer than the Asian group, although it is worth noting (as Islam and Baggio do, p. 301) that Nepal and Bangladesh are usually considered as South Asian countries, while MMNS's (2004) hypothesis regards East Asians.

Let us wrap up some conclusions concerning the RI studies on the Jonah Case. The results in Islam and Baggio (2020) suggest that participants tend to endorse the causal-historical intuition. That trend is numerically less robust among Norwegians than among immigrants from Nepal and Bangladesh, and participants living in Bangladesh. As regards specifically the comparison between Americans and Chinese, one can argue that Beebe and Undercoffer (2015, 2016) successfully address certain potential issues afflicting the material in MMNS (2004), and thus scholars should attribute greater dependability to the trend emerging from their material. This conclusion finds additional support in the higher statistical power that Beebe and Undercoffer's studies reach, resulting from their larger participant samples. The overall picture, however, becomes more intricate once one considers the most recent RI data on the Jonah Case: Devitt and Porot (2018) gather 91% of causal-historical answers with American participants (as I show in detail in the forthcoming section, 1.5.2). This proportion of answers supporting CHT strongly diverges from of MMNS's (2004) and, to an even greater extent, Beebe and Undercoffer's (2015, 2016). The vignette that Devitt and Porot develop is relatively short, retains a single orthographic form for the name and posits reference failure rather than the designation of a fictional entity. In light of these improvements, their results arguably possess a higher level of dependability compared to those that MMNS and Beebe and Undercoffer gather with their American samples. Undoubtedly, these varying outcomes indicate the need for further investigation in the domain of RIs concerning the

Jonah Case. In the forthcoming section, I also delve deeper into Devitt and Porot's findings, contextualizing their RI data within the broader scope of their research, which encompasses linguistic-usage tests as an integral and innovative component.

1.5.2 Linguistic-usage studies

Two studies explore the Jonah Case by employing TVJs and EP tests, which are supposed to unveil linguistic usage: Pinillos (2015) and Devitt and Porot (2018). Pinillos (2015) presents the following vignette to a group of American participants:

> The medievals used to believe in a goblin named Benjaminus. Benjaminus was widely thought to fly around the night sky and also thought to have the power to read people's minds. According to legend, Benjaminus liked to scare the local population by going in their houses and moving objects around. Of course, we know now that nothing in fact ever met the description associated with "Benjaminus". However, historians in Germany have discovered that the tales of Benjaminus were based on a real person, Benjamin Smith, who was a skillful night thief in medieval London. (Pinillos, 2015, p. 149)

Participants are then asked to express their level of agreement with two sentences:

(a) Benjaminus existed,
(b) Benjaminus is Benjamin Smith / (c) Benjamin Smith is Benjaminus

Participants use a scale of agreement ranging from 1 to 5, where "1" is "Strongly disagree", "5" is "Strongly agree" and "3" represents the neutral midpoint. Pinillos takes CHT to imply that "Benjaminus" refers to Benjaminus Smith, as Benjamin Smith would be the individual to whom the chain of communication behind "Benjaminus" leads. DT would entail that "Benjaminus" fails to refer, as no entity satisfies the description associated with the name, i.e. being a goblin flying around the night sky. Based on that, Pinillos expects participants endorsing CHT to agree with both sentences, while the ones endorsing DT to disagree.

Participants express an average mean of 2.13 with (a). The means for (b) and (c) do not differ significantly, and therefore Pinillos aggregates them. The mean for (b) / (c) is

3.41. The mean for (a) differ statistically significantly from the midpoint, while the mean for (b) and (c) does not. Moreover, the mean for (a) differs statistically significantly from the mean for (b) / (c). In a word, participants tend to disagree with (a) and are overall neutral with (b) / (c). Based on these results, Pinillos concludes that neither DT nor CHT is the correct theory of reference.

One notable issue with this vignette is that it can incorporate a case "fictional incorporation" (Johnson and Deutsch, 2021). Over time, some speakers may intentionally choose to use a name to talk about a fictional entity rather than the original real one: although a real individual inspires the fictional character, they are distinct entities. This is a phenomenon of "fictional incorporation" because speakers incorporate the name within a fictional context to talk about an entity derived from reality but divorced from it. If that is the case with "Benjaminus", then some speakers started using the name *without the same intention* as the people from whom they learned it. Therefore, one of the conditions for the continuity of the causal-historical chain collapses. Kripke (1980) points out:

> When the name is passed from link to link the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it. If I hear the name "Napoleon" and decide it would be a nice name for my pet aardvark, I do not satisfy this condition.
> (Kripke, 1980, p. 96)

As a consequence, also CHT would predict that "Benjaminus" refers to a person who does not really exist rather than the original bearer of the name. Thus, the vignette would fail to differentiate between the two theories, as the predictions of DT and CHT coincide. Several factors contribute to the interpretation of the vignette as an instance of fictional incorporation.

First, the vignette employs two names: "Benjamin Smith" and "Benjaminus". A participant may engage in the following reasoning: some speakers, at a certain juncture, made the deliberate choice to discontinue talking about the real entity, in order to talk instead about a fictional one. To signal this intention, they introduced a similar yet distinct name, that is "Benjaminus" as opposed to "Benjamin Smith".

The second aspect supporting the interpretation of this vignette as an instance of fictional incorporation pertains to the descriptions attributed to "Benjaminus", which delineate a prototypical fictional entity, specifically a goblin. This reinforces the construal that, at a specific juncture, certain speakers decided to cease denoting the original entity and instead engaged in a myth-creating endeavor inspired by reality, and yet distinct from it. Thus, a new fictional entity emerges, determining the disruption of the extant causal-historical chain and the appearance of a new one. Both DT and CHT would imply that the name refers to this new fictional entity. It is worth noting that the phenomenon of fictional incorporation need not exclusively involve supernatural entities such as goblins, ghosts, or wizards. Speakers may have chosen to use the name "Benjaminus" to talk about a fictional human being devoid of supernatural abilities. Nonetheless, the explicit appeal to supernatural creatures and attributes that are present only in imaginary narratives may suggest more straightforwardly that the vignette portrays a case of fictional incorporation.[53]

Third, the Jonah Case in MMNS (2004) presents the name "Attila" as the outcome of a series of small, and presumably unintentional, alterations of "Raditra". Therefore, it becomes less plausible for a participant to interpret "Attila" as a name that some speakers deliberately introduced to signal their intention to talk about a new fictional entity. In contrast, Pinillos' vignette merely states that the tale about Benjaminus were «based upon» the real individual. The phrase "based upon" exhibits a certain level of ambiguity, admitting the possibility that some speakers derived inspiration from the real individual to construct a fictional character.

One may object against the criticism of fictional incorporation by noting that the vignette is set in Medieval times, a period when people believed in supernatural entities such as goblins and ghosts. Such a context may disincline participants to interpret "Benjaminus" as undergoing a process of fictional incorporation. While this objection may carry some weight, it is still arguable that vignettes generally benefit from avoiding prototypically magical characters like goblins.

Another potentially significant problem lies in the vignette's susceptibility to the New Meaning Objection. Pinillos present (a), (b) and (c) without attributing them to a

[53] Kripke himself considers the possibility that the incorporation of mythical traits may break a causal-historical chain (2013, p. 50).

hypothetical speaker. Therefore, the descriptivist participant may arguably interpret the sentences based on the description that she herself associates with "Benjaminus". This description arguably reflects the complete information that the vignette provides and reads as follows: "The skillful night thief in medieval London who originated certain goblin stories". Hence, the descriptivist participant too will use the name to talk about Benjamin Smith.

At this stage, one may wonder what explains the discrepancy between the results for sentences (a) and (b)/(c). While this question is undoubtedly interesting, it lies beyond the scope of my argument. My primary interest is to show that the vignette and the final prompt that Pinillos employs suffers from some flaws and thus may not provide reliable results to test theories on the reference of proper names.

As already mentioned, Devitt and Porot (2018) undertake RI, TVJ and EP tests similar to those conducted for their Gödel Case, but in the context of a Jonah-style vignette. They use the following scenario:

> In Belgium, every schoolchild is told about a man named "Ambiorix" in history classes. They are told that he led the Gauls against the Romans when the Romans invaded Gaul, and that he killed scores of legionnaires single-handedly. These tales are all that is on the records today under "Ambiorix". Recently, it's been discovered that those tales are just legends. The leader known in ancient times as "Ambiorix" never fought the Romans, and he never killed a single legionnaire. The discovery reveals, however, that he was a much admired general and an important peacetime leader. He also had a charismatic personality, attracting the attention of those around him wherever he went. In ancient Gaul, leaders were often respected for their accomplishments on the battlefield. So his friends and fellow soldiers, who admired him greatly, spun those tales about him.
> (Devitt and Porot, 2018, p. 1563)

Stated below are the final prompts:

> 1. EP
> In light of this discovery, please say what if anything you think Belgium history books should now say about Ambiorix. Please give reasons but be brief.
> (Devitt and Porot, 2018, p. 1566)

2. TVJ-DT

Taking the preceding passage to be historically accurate, please say whether the following statement is true or false.

"Ambiorix did not really exist"

(*Ibidem*)

3. TVJ-CHT

Taking the preceding passage to be historically accurate, please say whether the following statement is true or false.

"Ambiorix was a real historical figure about whom stories have grown"

(*Ibidem*)

4. RI

In light of this discovery, who has the name "Ambiorix" been referring to in history classes?

(a) The man who those legends were first spun about.

(b) No real person at all

(*Ivi*, p. 1567)

In contrast to the majority of the previously examined Jonah cases, the vignette and prompts that Devitt and Porot conceive deviate in at least three key aspects, which contribute to a more dependable design. First, the vignette is relatively short. Second, the vignette uses a single orthographic form, "Ambiorix", instead of a name undergoing a process of alteration. Third, the final prompts abstain from making any reference to fictional entities and instead employ a referential-failure language.

Table 1.6 reports the percentages of causal-historical answers across the various tests.

Table 1.6 Results of the Jonah Case in Devitt and Porot (2018)

Prompt	**Percentage of causal-historical answers**
EP	100
TVJ-DT	81
TVJ-CHT	96
RI	92

All percentages are strongly causal-historical. Similarly to the Gödel Case, the most innovative part with the Jonah case that Devitt and Porot test is the EP one. The results are remarkable, with *all* participants (100%) expressing a judgment to the effect that "Ambiorix" refers to the original bearer of the name. However, analogously to what happens with the Gödel-style scenario, a critic may potentially consider the EP tests as susceptible to the New Meaning Objection: the participant reads the entire vignette and thus associates the name "Ambiorix" with a description reflecting all the information that the narrator provides. The description is: "The peacetime leader about whom some of his friends spun glorious-warrior stories". Based on this description, the participant proceeds to perform the EP tests, and the predictions from DT and CHT converge. An analogous criticism extends also to the TVJ tests, but probably not to the RI one. Unlike the other tests, the RI test indicates the speaker(s) using the name: «In light of this discovery, who has the name "Ambiorix" been referring to *in history classes*?» (my emphasis). The phrase «in history classes» specifies that the question regards the participant's referential intuition about the name *as* Hong Kong's teachers and students use it. As mentioned, Devitt and Porot's strongly causal-historical proportion of RIs, amounting to 92%, diverges from the results of the previous literature, thereby indicating the need for further research about RIs.

Unlike what they do with their Tsu Ch'ung Chih scenario, Devitt and Porot do not conduct a follow-up study for their Ambiorix Case EP test. Although Devitt and Porot

acknowledge that the New meaning Objection might apply to the Gödel Case, albeit implausibly, they consider that objection as a non-starter in the context of the Jonah Case. In the next chapter (Chapter 2), I will delve into their reasons.

2 A critical analysis of the linguistic-usage studies

Martí and Devitt are right that scholars should test theories of reference against linguistic usage and not referential intuitions. It is important to note that this approach does not advocate for a complete detachment between the studies on linguistic usage and those on referential intuitions. Rather, the experimental exploration of referential intuitions can serve the purpose of identifying specific challenges that warrant attention also within the investigation on linguistic usage. For example, the vignette preceding a RI final prompt often exhibits similarities with the one preceding a TVJ or EP prompt; therefore, the challenges related to the formulation of the vignettes tend to be analogous. Moreover, issues pertaining to a RI final prompt may also apply to TVJ and EP ones. For instance, as illustrated in section 1.3.2, Domaneschi and Vignolo (2020) argue that the epistemic ambiguity that they uncover in TVJs is an extension of the one that Sytsma and Livengood (2011) identify in RIs.

This chapter evaluates the adequacy of the linguistic-usage tests that scholars have employed in the literature thus far, namely EP and TVJs. I critically analyze the New Meaning Objection against Devitt and Porot's (2018) studies. In contrast to Devitt and Porot, I argue that the New Meaning Objection, once one frames it in terms of an *alternative* rather than a *new* descriptivist meaning, is theoretically well-founded and constitutes a real threat to the EP tests. I also argue that Devitt and Porot's follow-up experiment addresses the objection with a carefully crafted and effective methodology. Turning to the TVJ tests, I discuss Domaneschi and Vignolo (2020) and Li (2021), as they represent the most recent and comprehensive TVJ experiments available. While these studies exhibit a commendable sensitivity to the epistemic ambiguity inherent in TVJs, they suffer from other issues that impinge upon their reliability.

2.1 Devitt and Porot (2018)

For easy reference, let us revisit the EP final prompts in Devitt and Porot (2018). In the Gödel Case, it is:

> Having read the above story and accepting that it is true, what is your opinion of Tsu Ch'ung Chih?
> (Devitt and Porot, 2018, p. 1565)

In the Jonah Case, the prompt is:

> In light of this discovery, please say what if anything you think Belgium history books should now say about Ambiorix. Please give reasons but be brief.
> (Devitt and Porot, 2018, p. 1566)

For the sake of simplicity, the forthcoming analysis focuses on the Gödel Case. However, I take it to apply to the Jonah Case as well, although I spell out some specific considerations on the Jonah Case in section 2.1.1.5. According to the New Meaning Objection, an EP final prompt is flawed because the vignette introduces a new meaning of "Tsu Ch'ung Chih" for the descriptivist participant. Thus, two meanings are available. One pertains to the Hong Kong linguistic community, according to which "Tsu Ch'ung Chih" refers to the discoverer of the scientific result. The other meaning, the newfound one, is in force within the community of people "in the know", namely the speakers who have complete access to the events and associate "Tsu Ch'ung Chih" with a description picking out the thief. The latter meaning emerges as a consequence of the privileged information that the vignette conveys, leading the descriptivist participant to perceive the vignette's phrasing as establishing a new meaning. Therefore, the descriptivist gains access to two reference-fixing descriptions: not only the one that Hong Kongers share, but also the one reflecting the full information that the narrator provides. The latter description reads as follows: "The person who stole the result and claimed credit for it". If a descriptivist participant employs the name according to this reference-fixing description, the predictions of DT and CHT coincide. This flaw renders the EP prompt inadequate for effectively testing DT and CHT.

Devitt and Porot claim that the New Meaning Objection is misconceived. They advance three reasons. First, although logically possible, it is implausible that a new descriptivist meaning emerges. Second, if such a meaning were to arise, participants should exhibit some degree of oscillation between the two meanings in their answers; yet, this pattern of vacillation does not emerge. Third, the follow-up experiment yields consistent results when compared to the original study. In the forthcoming sections, I critically analyze these three criticisms.

2.1.1 First criticism: does a new meaning emerge?

2.1.1.1 The vignettes prior to Devitt and Porot's (2018)

As mentioned, Devitt and Porot's first criticism against the New meaning Objection is that there is no reason to posit that – assuming the truth of DT – their vignettes introduce a new meaning, different from the one that Hong Kongers share. To analyze Devitt and Porot's claim, it is necessary to take a step back and develop some considerations regarding the vignettes that scholars used in the previous studies, as the phrasing that Devitt and Porot employ for their scenarios stems from some critical reflections on the pre-existing experimental literature. Under the assumption that DT is the correct theory of reference, the narrator of the vignettes preceding Devitt and Porot's investigation does use the name according to a reference-fixing description that diverges from the one that is in force within the vignette's linguistic community (i.e., the community without privileged access to the events). Let us consider the original vignette in MMNS (2004):

> Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called "Schmidt", whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus, he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name "Gödel" are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel.
> (MMNS, 2004, p. B6)

When MMNS write «Now suppose that Gödel was not the author of this theorem», they use the name to talk about the thief. MMNS invite participants to imagine that in the vignette's fictional scenario a thief stole the theorem from another man, and they do so by employing "Gödel" as a name of the former individual. The same use emerges in the following sentence: «His friend Gödel somehow got hold of the manuscript and claimed

credit for the work, which was thereafter attributed to Gödel». Also in this context, "his" refers to the real discoverer of the theorem, while "Gödel" to the thief. Therefore, regardless of the correct theory of reference, it is beyond doubt that, in the narrator's language, "Gödel" is a name of the thief. Under the assumption that DT is the correct theory, within John's linguistic community "Gödel" is a name of the real discoverer of the theorem, via the description D_1 "The author of the incompleteness theorem". Instead, for the narrator and the people with complete epistemic access to the events, "Gödel" is a name of the thief, via the description D_2 "The man who stole the theorem and claimed credit for it". DT can admit this discrepancy, as it is possible for the same word, in this case a proper name, to have different referents in distinct linguistic communities, owing to the diverse reference-fixing descriptions.

Although Devitt and Porot call the criticism "New Meaning Objection", in vignettes like MMNS's (2004), D_2 seems to be an alternative descriptivist meaning rather than a new one. In other words, MMNS's vignette does not generate a new descriptivist meaning instantaneously; instead, it discloses that there are two pre-existing and equally valid descriptivist meanings, each selecting a distinct referent through its own reference-fixing description. Within John's linguistic community, "Gödel" has an established reference-fixing description, D_1, which picks out the real discoverer. Conversely, within the narrator's linguistic community, including the narrator herself and the people having privileged access to the events (as the participants), the speakers associate the name with another equally established reference-fixing description, D_2, which selects the thief. In light of these reasons, the label "alternative meaning" more crucially targets the real potential issue with the EP test. If that is correct, vignettes like the one in MMNS's (2004) present the descriptivist participant also with a meaning that deviates from the one in force within John's linguistic community. Therefore, the participant has the option to adopt the alternative meaning; if that is the case, her answer becomes inadequate for comparing DT and CHT, since both theories predict that the participant uses the name to talk about the thief.

The above conclusion applies to all the vignettes in the existing literature, including those that Devitt and Porot formulate. However, as previously mentioned, Devitt and Porot's scenarios introduce a significant element of novelty that stems precisely from a critical analysis of vignettes such as MMNS's (2004). Unlike the previous experiments,

Devitt and Porot's narrator never *uses* the critical name, but just *mentions* it between scare quotes. For example, in the Tsu Ch'ung Chih Case, the narrator states that «*a man called "Tsu Ch'ung Chih"* first determined the precise time of the summer and winter solstices», and not that "Tsu Ch'ung Chih first determined…" (p. 1562, my emphasis). The rationale behind that choice and its potential implications for the New Meaning Objection are the central focus of the forthcoming sections

2.1.1.2 The purported inconsistency of MMNS's (2004) vignette

Building upon Devitt's earlier work (2012b), Devitt and Porot choose to have the narrator *mention* the critical name, rather than use it, because a vignette like the one in MMNS (2004), in which the narrator employs the critical name (e.g., "Gödel"), would incorporate a bias against DT. Devitt and Porot argue that MMNS's (2004) vignette portrays a textual composition that a descriptivist narrator could not utter or write, thereby providing «unwittingly (and ironically)» evidence against DT (Devitt, 2012b, p. 27). Devitt wonders: according to DT, to whom do MMNS ordinarily refer when using "Gödel" in their philosophy classes?

> There can be no doubt that these philosophers [MMNS] are fully competent with the name "Gödel". And the referent of this name out of the mouths of the fully competent is the eminent logician who did in fact prove the incompleteness of arithmetic and spent many years at Princeton.
> (Devitt, 2012b, p. 27)

MMNS talk about «the eminent logician» and, according to DT, the description "The discoverer of the incompleteness theorem" mediates the reference. However, at approximately half the length of the vignette, MMNS's text reads: «Suppose that Gödel was not the author of this theorem». This passage is problematic, according to Devitt. Let us assume that MMNS are descriptivist users of names and employ "Gödel" as a shorthand for "The discoverer of the incompleteness theorem". If that is the case, they could not be disposed to utter the sentence "Suppose that Gödel was not the author of this theorem", as MMNS «in one and the same breath, both refer to Gödel and suppose away the basis of that reference» (Devitt, 2012b, p. 28). To illustrate the point, Devitt invites the reader to consider a sentence like "Suppose that the bachelors in Iceland are married"

(*ibidem*). Devitt argues that this sentence verges on self-contradiction. If – as plausible – a descriptive semantics covers "bachelor", then this term is equivalent to "adult unmarried male", where being unmarried constitutes a defining property of the meaning of "bachelor". Therefore, a speaker cannot conceive a situation or a country, like Iceland, in which bachelors are married, because this supposition denies a property by virtue of which a bachelor is such. For analogous reasons, if DT is correct and "Gödel" is equivalent to "The discoverer of the incompleteness theorem", a descriptivist narrator could not say "Let us suppose that Gödel did not discover the incompleteness of arithmetic". Therefore, Devitt and Porot conclude that MMNS's (2004) vignette stands as evidence against DT or, at the very least, is less than optimal, since it incorporates a bias against one of the two theories at test.

2.1.1.3 Strong Descriptivism and Weak Descriptivism

If Devitt and Porot's arguments are successful, they demonstrate that MMNS's use of the name "Gödel" is inconsistent with what we can call "Strong Descriptivism", according to which the meaning of a name is a definite description. The description itself is the name's semantic value and is part of the proposition that the sentence containing the name expresses. However, this is not the only way to understand descriptivism. According to what we can call "Weak Descriptivism", the description does not provide the meaning of the name but fixes its reference. The referent, which is constant across all possible worlds, is the name's semantic value and is part of the proposition that the sentence containing the name expresses. The labels "Strong Descriptivism" and "Weak Descriptivism" derive from Ichikawa, Maitra and Weatherson (2012), who consider the former version of the theory "stronger" because it assigns higher commitments to descriptions, which do not only fix the reference of names but also give their meaning.

Contrary to what Devitt and Porot assume, Strong Descriptivism is not necessarily the version of the theory that holds the utmost relevance to the ongoing debate. The primary interest of MMNS (2004) is the comparison of theories of *reference*. While the first sentence in MMNS (2004) states that «theories of *meaning* and *reference* have been at the heart of analytic philosophy since the beginning of the twentieth century» (p. B1, my emphasis), throughout the article MMNS only talk of "theories of reference" and the word "meaning" never occurs again. MMNS's focus becomes totally explicit when MSD

(2015) state that «the topic of MMNS (2004) is not […] "meaning descriptivism", but rather the "reference descriptivism"», which centers around the question of «how the *reference* of a proper name is determined» (p. 64, n. 5, emphasis in the original; see also MMNS, 2013, pp. 620-1). MMNS do not endorse Weak Descriptivism, which *denies* that the description provides the name's meaning. Instead, following their terminology, MMNS endorse "Reference Descriptivism": they assert that the description fixes the name's reference, but *remain neutral* about the possibility that it also provides the name's meaning.[54] Stated differently, Reference Descriptivism is an ambivalent stance that a theorist can develop both into Weak and Strong Descriptivism. Thus, MMNS's Reference Descriptivism is compatible with the view that Weak Descriptivism embodies, according to which "Gödel" refers to the individual X who discovered the theorem in the actual world, X constitutes the name's entire semantic value and X remains the constant referent of "Gödel" across all the possible worlds in which X exists.[55] The name refers to X also in those possible worlds, different from the actual one, in which X did not discover the theorem. Hence, MMNS do not contradict their theoretical assumptions when they suppose that Gödel did not discover the theorem and portray him as a thief. Philosophically speaking, MMNS outline a possible world in which Gödel is not the discoverer of the theorem. Weak Descriptivism allows for the conceivability of such a world.

My second criticism is that MMNS's formulation is not a *clear* piece of linguistic usage against Strong Descriptivism either. Plausibly, "The discoverer of the incompleteness of arithmetic" is not the only description that MMNS associate with "Gödel". Given their philosophical preparation, MMNS associate "Gödel" also with the discovery of the completeness of first-order logic and the formulation of the consistency-of-continuum hypothesis, among other things. According to Searle's version of DT, the name's referent is the *best* satisfier of that set of descriptions. Therefore, the referent of "Gödel" across possible worlds does not have to satisfy all the descriptions that MMNS associate with the name. As long as Gödel retains most of the properties different from

[54] MSD (2015) do not deny that the question of whether descriptions may also provide the meaning of names is «an interesting venue for future research» (p. 64, n. 5).

[55] "Gödel" fails to refer to any individual in those possible worlds in which X does not exist. Naturally, that does not mean to deny that the languages that people speak in those possible worlds may encompass "Gödel" as a name that successfully refers to some individual different from X. The conceptual paradigm underlying my analysis is one in which we evaluate the referent of "Gödel" across possible worlds in light of the reference-fixing mechanism that the name acquires within the language that people speak in the actual world.

being the discoverer of the incompleteness theorem, MMNS's counterfactual situation in which Gödel did not prove the incompleteness theorem is conceivable and consistent. Hence, MMNS's vignette does not provide clear-cut evidence against Strong Descriptivism either.[56]

It is worth emphasizing that descriptivism comes in many guises. For example, Strong Descriptivism could accommodate MMNS's supposition that Gödel did not discover the incompleteness theorem even if the only description that MMNS associate with "Gödel" were "The discoverer of the incompleteness theorem". Strong Descriptivism can achieve this outcome by considering the name as equivalent to a rigidified description – a view that Reimer (2004a) supports.[57] Such a description encompasses an operator to the effect that, given a specific possible world, the description picks out in that world the same individual as the one who, in the actual world, discovered the theorem.[58] Typically, the adverb "actually" signals this operator. Therefore, the rigidified version of the description would be "The individual who *actually* discovered the incompleteness theorem", which always selects the same person, even in those possible worlds in which he/she did not discover the incompleteness theorem. If the name "Gödel" is equivalent to that description, a descriptivist participant can conceive a scenario in which "Gödel did not discover the theorem". One potential problem concerning rigidified descriptions lies in their failure to capture the natural reply that an ordinary speaker would provide if one asked her "Who is N"? If a speaker were to answer "Who is Gödel?", she would claim that "He is person who discovered of the incompleteness theorem", rather than "He is the person who actually discovered the incompleteness theorem", where "actually" means "in the actual world". One of the purported qualities of descriptivism is its ability to

[56] One of the challenges related to the thesis that the name's referent is the satisfier of *most* of the pieces of information that speakers associate with the name is the absence of a clear principle for identifying which attributes are reference-fixing. It is possible to argue that not all the information that speakers associate with a name is reference-fixing. Suppose for example that most speakers associate "Aristotle" with the property of being blonde. Is that particular detail relevant when determining the name's referent, or is it totally negligible in light of apparently more crucial attributes, such as Aristotle's authorship of certain groundbreaking philosophical works? The guidelines that language users should follow in such circumstances are not clear. Devitt and Sterelny (1999) point out this problem while discussing DT and its lack of a «principled basis» for the selection of the pieces of information determining the name's reference-fixing description (p. 51).

[57] To clarify: Reimer does not claim that all names are descriptive; she maintains that, if a name is descriptive, then it is equivalent to a rigidified description. The aim of the experimental research on the reference of proper names is to ascertain the extent to which our language comprises descriptive names, beyond the ones that philosophers commonly acknowledge as such, like "Jack the Ripper", "Vulcan" or "Julius" – the latter introduced by Evans (1979) to refer to whoever invented the zipper. See also Jeshion (2004) for the discussion of names, like those referring to future entities or certain mathematical objects, that are arguably descriptive and yet did not attract as much philosophical scrutiny as the above-mentioned classical instances.

[58] I am restricting my analysis to the possible worlds in which that person exists in the first place.

capture the explicit cognitive content that language users link to names. However, ordinary speakers arguably do not grasp the difference between actual and possible worlds consciously, and any inclusion of an operator like "actually" in the descriptions that they associate with names seems only implicit, at best.[59]

The disentanglement of all the versions of descriptivism exceeds this dissertation's goals. The construal of descriptivism at the center of the thesis will be Reference Descriptivism, which asserts that the name's description fixes its reference, while remaining neutral about whether it also provides its meaning. This choice arises not only from MMNS's endorsement of that interpretation, but also from the possibility of articulating Reference Descriptivism both into Weak Descriptivism and a rigidified version of Strong Descriptivism. A descriptivist seems to need one of these two theories anyway, as they are the only versions of descriptivism that can accommodate the verbal situation in which, given a name N and a (cluster of) reference-fixing description(s) D, a speaker supposes that N is not D. As argued, MMNS's supposition that Gödel did not discover the theorem can be compatible with Strong Descriptivism by virtue of the other pieces of information, different from the discovery of the incompleteness theorem, that MMNS associate with "Gödel". However, it seems linguistically possible for MMNS to suppose that Gödel neither discovered any of the mathematical achievements that they attribute to him nor engaged with any of the activities that they associate with that name. As Kripke (1980, pp. 61-3) stresses in his discussion of Searle's (1958) descriptivist proposal, a traditional version of Strong Descriptivism cannot accommodate that supposition, because, to use Devitt's words, MMNS would «suppose away the basis of [their] reference» (Devitt, 2012b, p. 28). If a descriptivist theorist wishes to allow for the formulation of such suppositions, Weak Descriptivism or a rigidified version of Strong Descriptivism seem the only viable alternatives. In other words, there seem to be reasons that make these two versions of DT preferable. Therefore, the possibility to elaborate

[59] It is possible to frame the variants of DT also in terms of their semantic and metasemantic commitments, as Häggqvist and Wikforss (2015) emphasize. Given a name N and a theory, the semantic commitment of the theory pertains to the semantic value attributed to N: the semantic value is the element that N contributes to the proposition that the sentence containing N expresses. Instead, the metasemantics concerns «*the facts in virtue of which* a term has a certain semantic value» (p. 111, emphasis in the original). As a way of example, let us consider Weak and Strong Descriptivism. The two theories diverge from a semantic point of view: according to Weak Descriptivism, N's semantic value is the object itself to which N refers; according to Strong Descriptivism, N's semantic value is the description with which N is associated. However, the two theories exhibit an analogous metasemantics: according to Weak Descriptivism, what determines N's referent is the description associated with it; according to Strong Descriptivism as well, the description is what determines N's semantic value, since the description itself is N's semantic value – ultimately, in Strong Descriptivism, semantics and metasemantics coincide.

descriptivism in the direction of these two theories seems to be a charitable move in support of the descriptivist. If even under this generous approach DT fails to meet the experimental challenge, *a fortiori* the empirical data provide conclusive evidence against any alternative version of descriptivism. In light of these reasons, in this dissertation "DT" labels Reference Descriptivism. While I call the objection at issue the "New Meaning Objection" to follow Devitt and Porot's wording, I henceforth discuss different "reference-fixing descriptions" rather than "meanings". This choice reflects my construal of the descriptivist view, which is neutral about the description's import for the name's meaning.[60]

To provide a comprehensive overview, it is finally worth mentioning at least another non-classical articulation of the descriptivist theory, known as "causal descriptivism" (Kroon, 1987). According to this construal of the theory, which Kripke himself considers in its broad outline (1980, p. 88 n. 38 and 162), the meaning of a name is equivalent to a description that outlines the name's bearer as the individual properly situated at the end of the name's causal-historical chain. For example, the meaning of the name "Gödel" that John utters would be "The individual at the end of causal-historical chain culminating in John's current use of 'Gödel'". Clearly, this theory is capable of accommodating any evidence that seems to support CHT: whatever the causal-historical referent is, that entity would automatically satisfy the pertinent causal description. However, such a construal of descriptivism exposes to several problems.

In the first place, causal descriptions do not identify the cognitive content that ordinary speakers associate with names (Kripke, 1980, p. 162; Pinillos, 2015, p. 142; Martí, 2017, 360-2). If asked for example who Gödel is, an ordinary speaker would not reply that Gödel is the individual who is properly situated at the end of the name's causal-historical

[60] Reimer underscores the necessity to accommodate suppositions like "Let's imagine that N is not D" even with descriptive names (2004a, p. 614). She considers what philosophers of language regard as almost the prototypical example of a descriptive name, that is "Jack the Ripper", following Kripke's (1980) discussion of it (pp. 79-80, 84). It seems highly plausible that the reference of "Jack the Ripper" works through a description such as the «serial killer active in and around the impoverished Whitechapel district of London, England, in 1888» (as Wikipedia's article page on Jack the Ripper reports: <https://en.wikipedia.org/wiki/Jack_the_Ripper>, consulted on December 18th, 2023). Reimer stresses that it is possible to conceive a situation in which Jack the Ripper does not satisfy that description, for example when one utters a sentence like "Suppose that Jack the Ripper's parents and friends had loved him; in such a circumstance, he would not have developed his psychopathic tendencies and would not have committed any murders". While Kripke himself seems to endorse the conceivability of such suppositions involving descriptive names (1980, p. 79), Devitt raises some doubts (2015a, pp. 136-7) and maintains that descriptive names, such as "Jack the Ripper", fail to clearly pass the standard tests of rigidity. If "Jack the Ripper" referred to the same individual across all possible worlds, a sentence like "Jack the Ripper might not have been Jack the Ripper" should be false and a sentence like "It might have been the case that Jack the Ripper was not a murderer" should be true. However, Devitt claims, this «seems not to be» (p. 137).

chain. Causal descriptivism shares this flaw with the rigidified version of descriptivism, and probably to a greater extent: while one might argue that speakers implicitly integrate an operator like "actually" into their reference-fixing descriptions, it is highly implausible that they possess the capacity to articulate such strongly theoretical descriptions as the causal ones.

Second, as Devitt and Sterelny emphasize (1999, p. 61), causal descriptivism expects that laypeople are able to articulate the distinctive features of the chain that make an individual properly situated at the end of it. However, even theorists have not fully accomplished that task: think for example of complex cases like "Madagascar".[61] Furthermore, the concept of causal-historical chain, along with its potential semantic implications pertaining to the reference of names, was not even at the center of philosophical circles prior to Kripke's groundbreaking *Naming and Necessity*.[62]

Finally, the most significant issue with causal descriptivism is its parasitic reliance on CHT: it postulates descriptions that center around causal-historical chains. There seems to be no point in positing these descriptions. Causal-historical chains themselves, plus the baptism act, seem capable of fulfilling all the theoretical work that theorists need to account for the referential connection between names and the extra-linguistic world. Causal descriptions seem superfluous in terms of explanatory value: they are «no more than [..] idle wheel[s]» (Martí, 2017, p. 362).[63]

[61] As Evans (1973) explains while introducing the term, the name "Madagascar" originally referred to a portion of mainland Africa, rather than the island that people currently know by that name. Marco Polo misunderstood how the locals were used to employing the name, and imported "Madagascar" into Europe as a name of the island. Kripke's (1980) proposal, which by his own admission is not a full-fledged theory but a «better picture» than the rival descriptivist view, entails that "Madagascar" still refers to the mainland portion (p. 93). As Kripke (1980) recognizes in his cursory treatment of the case within the *addenda* of his work (p. 163), the implication that "Madagascar" still refers to the mainland part of Africa is problematic. Therefore, a compelling solution is required. Devitt (1981, pp. 138-52) maintains that his more causally-marked version of CHT is able to account for cases like that. His idea is that, since Marco Polo was causally connected with the island and had it "in mind" – in Donnellan's (1966) sense – when using the name, he gave rise to a process that gradually resulted into a reference shift. For an analysis of the case, see Santambrogio (2014) and Fitch (2004, p. 67).

[62] On the *groundbreaking* character of Kripke's philosophy, including also his seminal work *Naming and Necessity*, see García-Carpintero (2023).

[63] For example, Devitt and Porot (2018) claim that their «results count strongly against classical description theories and hence give some *indirect* support to causal-historical theories of names» (p. 1555, my emphasis). The support for CHT is indirect because, *per se*, any experimental result compatible with CHT is also consistent with causal descriptivism. However, conceptual reflections show that causal descriptivism is not a viable alternative, thereby leaving CHT as the only theory capable of accounting for their results. It is possible to extend this conclusion to any experimental work whose results are compatible with descriptivism only under the adoption of causal descriptivism. To use Devitt's (2020, p. 437, n. 75) words: «Sterelny [2020] is quite right […] that these experiments [the experiments on the reference of proper names] do not discriminate "between causal descriptive theories and causal theories of reference" [p. 179, n. 6]. We don't need experiments to reject causal descriptivism». Whether any version of descriptivism, different from the causal one, is capable of accounting for Devitt and Porot's results is the object of discussion in this section.

2.1.1.4 Mention, use and alternative meanings

Notwithstanding the aforementioned concerns, Devitt and Porot consider MMNS's vignette and the other ones within the literature as presenting a text that a descriptivist narrator could not utter. To address this problem, Devitt and Porot intentionally introduce an innovative feature in their formulation. Therefore, the question of whether the New Meaning Objection applies to Devitt and Porot's material merits a specific analysis. Let us revisit the Tsu Ch'ung Chih Case:

> Students in astronomy classes in Hong Kong are told that a man called "Tsu Ch'ung Chih" first determined the precise time of the summer and winter solstices. This is the only thing that typical Hong Kongers ever hear about this man. Now suppose that that man did not make the discovery he is credited with. He stole it from an astronomer who died soon after making the discovery. But the theft remained entirely undetected and so the man that Hong Kongers have been told about became famous for the discovery of the precise times of the solstices.
> (Devitt and Porot, 2018, p. 1562)

Devitt and Porot stress that they never *use* the name "Tsu Ch'ung Chih", but just *mention* it by scare quotes (pp. 1562-3). Specifically, they mention the name at the beginning of the vignette, by talking of a «man called "Tsu Ch'ung Chih"». Throughout the text, they subsequently refer to that individual by the anaphoric expressions "he" or "that man". Hence, when Devitt and Porot write «suppose that that man did not make the discovery he is credited with», they do not assert "suppose that *Tsu Ch'ung Chih* did not discover the solstice times". Rather, they state "suppose that *a man called 'Tsu Ch'ung Chih'* did not discover the solstice times". In this way, Devitt and Porot avoid an assertion that, according to them, a narrator cannot utter if Strong Descriptivism is true.

Considering that the focus of this dissertation revolves around Reference Descriptivism, it is not primarily relevant whether Devitt and Porot's formulation is consistent with Strong Descriptivism. Instead, the key question is whether Devitt and Porot's formulation effectively tackles the New Meaning Objection. In other words, it is important to examine whether their choice of mentioning rather than using "Tsu Ch'ung Chih" directly affects the potential ambiguity between two reference-fixing

descriptions.[64] Devitt and Porot argue that it is implausible that the vignette generates a new reference-fixing description of "Tsu Ch'ung Chih", different from the one that Hong Kongers share:

> To suppose that the new associations are reference-determining is to suppose that the participants take the story narrated in a vignette to have led to a new conventional meaning and reference for the name; to have led to a community, including the vignette's narrator, that uses the name according to a new convention. Now meanings can change, of course, but they don't change every time we get new information about the world. In particular, a proper name typically does not change its meaning on discovery that it is "empty". Indeed, it was a triumph of the description theory that it gave a nice account of "Zeus does not exist", an account that presupposes that the "Zeus" means now just what it meant before it was discovered that nothing fits the descriptions associated with it. So we think it clear that the New-Meaning objection fails in the Ambiorix case: The description theory would not take the discovery in that case to have created a new meaning of "Ambiorix" and would predict, just as we supposed, that participants will take Ambiorix not to exist. And nothing in the Tsu Ch'ung Chih vignette invites the idea that "Tsu Ch'ung Chih" has developed a new meaning. So we think that a plausible description theory would not suppose that there was a new meaning there either.
> (Devitt and Porot, 2018, p. 1569)

At this juncture, the difference between the notion of a "new" descriptivist meaning and that of an "alternative" one becomes crucial, as anticipated at the beginning of this chapter. What I have argued so far is that vignettes like MMNS (2004) present an *already established* reference-fixing description, D_2, in force within the narrator's linguistic community and amounting to "The thief who appropriated the result and claimed credit for it". This description is *alternative* to another equally established reference-fixing description, D_1, in force within the fictional Hong Kong community and amounting to "The discoverer of the scientific result". Instead, the notion of a "new" meaning that Devitt and Porot put forth describes a *change* that one reference-fixing description, D_1,

[64] The concept of ambiguity traditionally pertains to *multiple meanings*. However, as explained above, my understanding of DT maintains a neutral stance on the hypothesis that descriptions also provide the meaning of names. Therefore, my use of the term "ambiguity" and the derived terminology is somewhat peculiar, since I apply it to alternative reference-fixing descriptions.

undergoes in light of the discovery of some facts, resulting into a new reference-fixing description, D_2. As a consequence of this alteration, Hong Kongers possess only the old reference-fixing description, D_1, while the narrator, the participant and the community "in the know" would also possess the "new" description, D_2.

Devitt and Porot are right that «meanings can change, of course, but they don't change every time we get new information about the world» (p. 1569). In other words, only an implausible version of DT can admit that a participant, subsequent to reading the vignette, interprets the privileged information therein as effecting a change of D_1 and promptly generating a new conventional reference-fixing description for the name, that is D_2. In the absence of an explicit baptizing event, modifications in the reference-fixing descriptions do not take place abruptly; rather, such changes can achieve a conventional status only once the practice of using the name according to the new description gradually becomes established. However, D_2 may raise issues for Devitt and Porot's experiment not because a descriptivist participant can likely take the vignette to *immediately* impose D_2 as a new conventional description. Rather, the issue lies in the participant's potential perception of D_2 as a *previously established* reference-fixing description in force within the linguistic community of people "in the know", which includes the narrator. In other words, the participant may consider D_2 as an already conventional reference-fixing description that has coexisted alongside the one that the Hong Kong community shares.

At this point, one may wonder how likely it is that a descriptivist participant takes D_2 as an established reference-fixing description. One may object that, since the narrator does not *use* "Tsu Ch'ung Chih", the vignette presents no linguistic evidence in which the name "Tsu Ch'ung Chih" refers to the thief via the description "The man who stole the astronomic result and claimed credit for it". Therefore, the participant would lack linguistic-usage grounds to take the above description as fixing the name's reference within the narrator's community. Such an objection is disputable, though. Devitt and Porot's choice to use "A man called 'Tsu Ch'ung Chih'", instead of directly employing "Tsu Ch'ung Chih", does not prevent the descriptive participant from supposing that the narrator is part of a larger community of people "in the know" where D_2 is the conventional reference-fixing description of the name.

Suppose that DT is the correct theory for names and that a speaker, John, utters a sentence such as "Today the most famous rock band in Italy, called 'Måneskin', will give

a concert in Milan". John does not use the name "Måneskin'", as he does not say that "Måneskin will give a concert in Milan"; he merely mentions the name. The addressee, Smith, is not an expert on rock bands and his mental lexicon does not include the name "Måneskin" prior to this conversation. It seems plausible for Smith to consider "Måneskin" as the conventional name for the above rock group, through the description "The most famous rock band in Italy" that John's utterance provides. Indeed, both CHT and DT concur in asserting that Måneskin are called "Måneskin". The divergence between DT and CHT lies in the mechanism by virtue of which Måneskin are called "Måneskin": DT posits the fulfillment of a description; CHT posits the path of a causal-historical chain. For that reason, employing the phrase "The group called 'Måneskin'" does not prevent Smith from perceiving "Måneskin" as the group's conventional name. An analogous line of reasoning extends also to the Tsu Ch'ung Chih Case. The vignette describes a linguistic community, the Hong Kong one, that associates "Tsu Ch'ung Chih" only with D_1, "The discoverer of the scientific result". However, the vignette also provides some privileged information pertaining to the "man called 'Tsu Ch'ung Chih'" and involving a theft. The descriptivist participant can posit that this comprehensive information outlines the name's reference-fixing description within the narrator's community, similarly to how an addressee can suppose that the information that she just gained about the group called "Måneskin" – namely its being the most famous rock band in Italy – conventionally fixes this name's referent.

Devitt (personal communication) acknowledges that, since both theories – DT and CHT – agree on the fact that a person "is called" by a certain name, the descriptivist participant might potentially take the phrase "the person called 'Tsu Ch'ung Chih'" to indicate the conventional name by which the narrator's linguistic community calls the thief. Devitt further recognizes that, probably, a more explicit formulation such as "A man who was *then/at the time* called 'Tsu Ch'ung Chih'" might have been more effective in discouraging the descriptivist participant from taking the name as the one by which the people "in the know" *currently* refer to the thief. If my criticism is correct, Devitt and Porot's vignette provides the participant with two reference-fixing descriptions: one in force within the narrator's community and the other within the Hong Kong's.

2.1.1.5 The New Meaning Objection as applied to the Jonah Case

The preceding arguments revolve around the Gödel Case: however, they arguably extend to the Jonah Case as well. The descriptivist participant can assume that, for instance, the reference-fixing description for "Jonah" within the narrator's linguistic community diverges from the one within the community that the vignette describes, thus enabling the descriptivist participant to answer according to the purportedly causal-historical lines. As seen, in their critical examination of the New Meaning Objection, Devitt and Porot primarily rely on the illustrative example of the name "Zeus". They argue that it was a triumph for the descriptivist theory to provide an elegant analysis of the sentence "Zeus does not exist", as the name did not change its meaning as a consequence of the discovery of the non-existence of Greek gods. Given that "Zeus" is an empty name, Devitt and Porot use this example to show the clear invalidity of the New Meaning Objection when applied to the Ambiorix Case, namely a scenario revolving around a name that DT regards as empty. Instead, they take the objection to possibly apply to the Tsu Ch'ung Chih Case, albeit implausibly: for that reason, they conduct a follow-up experiment only in the context of their Gödel scenario.

In contrast to Devitt and Porot, I have argued that the problem marring their experimental scenarios concerns the possibility of an *alternative* reference-fixing description, rather than a *new* one. Therefore, their observation that "Zeus" did not undergo a change in its reference-fixing description following the discovery of the non-existence of Greek gods does not directly address my concern. Moreover, Devitt and Porot's conclusion regarding the inapplicability of the New Meaning Objection to the Jonah Case is also due to their questionable choice of using the example of "Zeus", which is somewhat different from the typical names at the center of the Jonah cases. Let us consider a possible experimental vignette revolving around the name "Zeus". It would describe a linguistic community, like the Ancient Greek one, that associates the name with a description such as "The king of Greek gods, who rules on Mount Olympus". Then, the narrator would state that our contemporary and well-informed linguistic community knows that the description associated with the name fails to pick out any real entity, because things such as gods launching their lightnings from the sky do not exist. Therefore, what the contemporary community knows, compared to the Ancient Greek one, is not that in the past "Zeus" was a name of some real individual whose actions

generated certain legendary stories, but rather that no entity has ever satisfied the description associated with the name. In other words, we became cognizant that "Zeus" has always been empty, that is, a name that has never referred to any real entity *from its inception*. To use Devitt's (1981) terminology, "Zeus" is a «failed name» (p. 175). Crucially, the Ambiorix Case outlines a different scenario. Whichever the right theory of reference is, it is beyond doubt that *the vignette presents a real individual whom his contemporaries called "Ambiorix"*, that is, the peacetime leader. The narrator explains that some tales originated from the actions that this person performed. In light of this, the descriptivist participant can combine the above pieces of information and posit that the name, within the community "in the know", refers to the individual satisfying the description "The peacetime leader whose actions made his friends generate some warrior tales". If that is correct, an alternative reference-fixing description is available and the name's referent is ambiguous. Therefore, the points that I have developed for the Tsu Ch'ung Chih scenario also extend to this context, and the Ambiorix Case does not evade the challenge that New Meaning Objection poses.

An important difference between the Tsu Ch'ung Chih Case and the Ambiorix Case emerges, though. In the latter, while providing the privileged information to the participant, the narrator presents the peacetime leader as "the leader known *in ancient times* as 'Ambiorix'" (p. 1563, my emphasis). As mentioned (section 2.1.1.4), phrases such as "The individual then / at the time / in ancient times called N", rather than the simpler "The individual called N", are arguably more effective in discouraging the descriptivist participant from taking N as the name by which the people "in the know" *currently* refer to the individual. By specifying that "Ambiorix" was the name by which the individual was known during a specific past era, the narrator arguably suggests to the descriptivist participant that this is not the name by which anyone, including the narrator's linguistic community, presently calls that individual. As a result, the only available reference-fixing description is the one that the vignette's Belgian community possesses. It is difficult to establish the precise extent to which the use of the phrase "known in ancient times as" effectively brings about the desired outcome. The fact that "Ambiorix" was the way people called the peacetime leader in the past does not necessarily prevent the descriptivist participant from supposing that this may *also* be the way a specific linguistic community – the narrator's – currently calls the same individual. That said, the

impact of the New Meaning Objection is admittedly more conjectural in the Ambiorix Case than in the Tsu Ch'ung Chih one. The New Meaning Objection may resurface if the descriptivist participant is not sufficiently sensitive to the phrase "known in ancient times" and thus takes the vignette as presenting the reference-fixing description in force within the narrator's linguistic community. In the forthcoming discussion, I take the New Meaning Objection to apply both to the Tsu Ch'ung Chih and the Ambiorix Case, although this Objection constitutes a threat against the latter scenario in a more speculative manner.

2.1.2 Second criticism: are there signs of ambiguity?

If the preceding arguments are correct, they show that the vignettes – under the assumption that DT is true – present two reference-fixing descriptions for the name. One description, D_1, is in force within the vignette's linguistic community; the other, D_2, is in force within the narrator's linguistic community, namely, the group of people "in the know". The participant has privileged information about the events and therefore is part of the narrator's linguistic community, gaining access to D_2. However, Devitt and Porot observe that, even accepting the above reconstruction, the descriptivist participant still has access to D_1 as well, as the vignette presents her with the reference-fixing description in force within the Hong Kong community (or the Belgian one, in the context of the Ambiorix Case). Therefore, when performing the EP test and expressing her opinion about Tsu Ch'ung Chih, the participant could choose to adopt D_1 and use the name according to it. Based on that observation, Devitt and Porot formulate their second criticism against the New Meaning Objection (p. 1569). A descriptivist participant, aware of the coexistence of the two reference-fixing descriptions, should show sensitivity to that ambiguity. Instead of exclusively providing negative responses such as "Tsu Ch'ung Chih is a thief and impostor", it seems reasonable to anticipate a range of answers in which the participants express positive judgments. After all, D_1 – the Hong Kongers' reference-fixing description – is accessible to the descriptivist participant and nothing seems to prevent her from expressing an evaluation based on it. Alternatively, one could expect participants to exhibit some edging answers, such as "Tsu Ch'ung Chih is a thief and impostor, but the person whom Hong Kongers know by that name is a great scientist". Crucially, Devitt and Porot do not gather answers of that nature. Therefore, they conclude

that DT fails to explain the results. It is beyond doubt that the name "Tsu Ch'ung Chih" has D_1 as its reference-fixing description among Hong Kongers. DT can account for the results if "Tsu Ch'ung Chih", alongside D_1, has another reference-fixing description, D_2, which is accessible to the narrator's linguistic community. Therefore, for the descriptivist participant, the name's referent is ambiguous between the satisfier of D_1 and the satisfier of D_2. However, the participants' answers do not show any sign of this ambiguity. Therefore, these responses seem to stem from participants who are not descriptivist, but rather causal-historical, as the Kripkean theory posits that "Tsu Ch'ung Chih" – in the context of the vignette – is not ambiguous.

One may question Devitt and Porot's reasoning by wondering whether a descriptivist participant should show such a sensitivity toward the Hong Kongers' reference-fixing description while formulating her EP answer. The fact that – according to DT – "Tsu Ch'ung Chih" is ambiguous does not necessarily imply that two descriptions hold equal significance. If the participant considers one description as more salient than the other, she might resolve the ambiguity by simply favoring the former. *Contra* Devitt and Porot, two reasons seem to suggest that the relevant description is the one to which the narrator's community has exclusive access.

First, the participant learns the story (and the name) from the narrator and therefore acquires an epistemically full window into the events, *thereby entering the narrator's privileged linguistic community*. It thus seems reasonable to expect that the participant, in her EP test, aligns her use of the name with the reference-fixing description, D_2, that she takes to be in force within the linguistic exchanges among the people of the narrator's community.

Second, the final prompt reads: "Having read the above story and accepting that it is true, what is your opinion of Tsu Ch'ung Chih?". This question uses the name "Tsu Ch'ung Chih"; moreover, the speaker asking this question is someone who is familiar with the story's content, as otherwise the final prompt could not mention it. Hence, the prompt originates from a person "in the know", someone who belongs to the narrator's linguistic community – maybe the narrator herself. If I am right that the descriptivist participant can reasonably take the speakers from the narrator's community to use "Tsu Ch'ung Chih" as a name of the thief, then the question asks an opinion about that person.

Therefore, the participant will express a negative opinion about that individual and will do so by employing the name that the final prompt itself uses, namely "Tsu Ch'ung Chih".

In conclusion, Devitt and Porot object that DT can account for the results only if the name "Tsu Ch'ung Chih" is ambiguous between two referents. Problematically, the participants' answers do not show any sensitivity to this purported ambiguity. In reply, I point out valid reasons to suspect that a descriptivist participant may take the description of the narrator's community to override the Hong Kongers'.

2.1.3 Third criticism: the follow-up experiment

Let us now examine the follow-up experiment, which constitutes Devitt and Porot's third reason for rejecting the New Meaning Objection. The final prompt undergoes a modification:

> Having read the above story and accepting that it is true, what would you say to a Hong Konger about Tsu Ch'ung Chih? (Please write as if you are speaking to a Hong Konger)
> (Devitt and Porot, 2018, p. 1573)

The follow-up final prompt still requires the participant's opinion about Tsu Ch'ung Chih; however, it does so by asking what the participant would say to a Hong Konger. Moreover, the parenthetical remark "Please write as if you are speaking to a Hong Konger" further stresses that the interlocutor belongs to the Hong Kong community. As argued, the participant belongs to the narrator's linguistic community and thus has access to the privileged reference-fixing description D_2. However, that does not change the fact that the participant also has access to the Hong Kongers' description, D_1, as she acquires it from the vignette. The rationale of Devitt and Porot's follow-up experiment is that pragmatic principles should guide the descriptivist participant to use the name according to the description available both to herself and to her interlocutor from Hong Kong, that is, D_1. Therefore, the New Meaning Objection cannot affect the follow-up final prompt, since the supposed alternative reference-fixing description D_2 is irrelevant in the conversational context.

Machery (2021) claims that the follow-up experiment does not address the New Meaning Objection successfully. Machery writes:

> […] in a follow-up study, [Devitt and Porot ask] participants to "give their opinions of Tsu Ch'ung Chih as if they were speaking to a Hong Konger". Because a Hong Konger in the fictional situation would not have access to the information given by the narrator of the vignette, Devitt and Porot assume that a descriptivist reader would only associate the description "the man who first determined the precise time of the summer and winter solstices" with "Tsu Ch'ung Chih". […] Their assumption is, however, questionable: in contrast to the speaker's intention to refer (which determines the speaker's reference), the description semantically associated with a proper name may not change as a function of the information available to a speaker's interlocutors.
> (Machery, 2021, p. 547)

Machery argues that the final prompt that Devitt and Porot outline may impact the participant's speaker-reference making, not her semantic-reference one. Therefore, if the participant adheres to the semantic referent of the term, she will give the same answer as the one that she would choose in the original experiment. If instead the participant chooses to use the name non-literally to make speaker reference, her answer will be unsuitable to compare DT and CHT, as these are theories on the names' semantic reference. Hence, the follow-up experiment would fail to successfully address the New Meaning Objection. Machery is right that the follow-up experiment assumes that the pragmatics of the conversation should lead the participant to favor one description of "Tsu Ch'ung Chih" over the other. However, I do not agree that this pragmatic constraint regards the choice of a non-literal reference-fixing description over a literal one. Rather, always under the assumption that DT is true, the constraint relates to the choice between two distinct conventions that the participant can use to formulate her answer. One is the convention in force within the narrator's community, where "Tsu Ch'ung Chih" literally refers to the thief; the other is the convention in force within the Hong Kong community, where "Tsu Ch'ung Chih" literally refers to the discoverer of the scientific result. The participant is part of the narrator's linguistic community, where the utterances containing "Tsu Ch'ung Chih" have the thief as the name's semantic reference. However, the

peculiar conversational context requires the descriptivist participant to temporarily abandon the convention in force within her community and import the one in force within the Hong Kong community. Once that shift happens, the participant deploys a conventional meaning according to which "Tsu Ch'ung Chih" literally refers to the thief, and thus that is the individual about whom she literally talks. In other words, the choice is between two different conventions with differing literal referents of "Tsu Ch'ung Chih"; once the participant makes the convention choice, the ensuing referent is the individual about whom the participant literally talks.[65]

More in general, Devitt and Porot's follow-up experiment presents an elegant and well-designed methodology for tackling the New Meaning Objection. Their results, where 89% of the answers are in line with the predictions deriving from CHT, constitute substantial anti-descriptivist evidence. That said, it is important to also investigate the other methodology that scholars employed to test linguistic usage, namely TVJs. As discussed, if production is one side of linguistic usage, comprehension is the other side, and TVJs shed light on it. Therefore, it is important to investigate the extent to which also that source of data provides results consistent with those that Devitt and Porot collect.[66]

2.2 Domaneschi and Vignolo (2020)

The forthcoming sections present a critical analysis of Domaneschi and Vignolo (2020) and Li (2021), which I take to be the most preeminent TVJ studies due to their sensitivity toward the so-called epistemic ambiguity. Both studies have the merit of identifying the possibility to approach the final prompt from the speaker's perspective, which is not suitable to compare the two classical theories of reference. Their results at least suggest that some participants in fact adopt that perspective during the experiment. However, I ultimately argue that these two studies raise legitimate concerns due to the theoretical nature of their follow-up questions.

[65] Note that the same phenomenon also takes place in the context of a TVJ test that specifies a speaker in semantic contradiction (e.g., Ivy). The participant belongs to the narrator's linguistic community, where the name "Tsu Ch'ung Chih" conventionally refers to the thief. However, given that Ivy is the person uttering the sentence, the descriptivist participant has to temporarily set aside the convention in force within her own community and evaluate Ivy's utterance in light of the convention in force within Ivy's community. That is the only convention relevant to judging the semantic referent that the name acquires within Ivy's sentence.

[66] Devitt and Porot (2018) themselves conduct TVJ tests. However, as seen (section 1.4.3.4), the absence of a speaker in semantic contradiction (e.g., John, Ivy, Emily and so on) makes those tests fall prey to the New Meaning Objection.

2.2.1 The study

The work by Domaneschi and Vignolo (2020) encompasses three separate experiments that they conduct under a between-subjects design, namely Experiment 1, Experiment 2, and Experiment 3. According to Domaneschi and Vignolo, these three studies show that the epistemic ambiguity afflicts the standard TVJ tests and lend support to CHT.

2.2.1.1 Experiment 1

Let us rehearse the main features of Experiment 1, presented in section 1.4.3.3. A group of Italian participants read a simplified version of the Super Dog Race, a Gödel-style vignette featured in Li et al. (2018).

> Long ago, there was a race called the Super Dog Race. Max, Pickles and Blaze participated in the race. Max crossed the finish line first, winning the race, but he got too excited and ran all the way to the North Pole. Pickles crossed the finish line second. He stopped and watched Max run away. The announcer of the race mistakenly thought that Pickles won the race. He told every newspaper in the world that Pickles won. He also told them that another dog, Blaze, ran very fast despite his short legs. Since then, everyone learned that Pickles won the race. They don't know anything else about Pickles.
> Tom and Emily learned at school that Pickles won the Super Dog Race. This is the only thing they know about the dog race and Pickles. They don't know anything about Max.
> (Domaneschi and Vignolo, 2020, pp. 448-9)

Participants perform a TVJ judgment test:

> That night, their dad asked: Do you know who won the Super Dog Race?
> Emily said: "Pickles was the dog that won the Super Dog Race".
> Do you think that her claim is
> (a) True or
> (b) False?
> (Domaneschi and Vignolo, 2020, p. 449)

The participants choosing the TVJ "True", namely the seemingly descriptivist answer, receive a follow-up question. Domaneschi and Vignolo aim to determine the epistemic perspective that they employed when formulating their TVJ.

> You think Emily's claim is true because:
> (a) Pickles did not win the Super Dog Race, but Emily believes that Pickles won the race because her teacher told it to her.
> (b) What Emily believes is a true description of Max and she uses the name "Pickles" to talk about the dog that really (unbeknown to Tom and Emily's teacher) won the Super Dog Race.
> (Domaneschi and Vignolo, 2020, p. 449)

Option (a) reflects the adoption of the speaker's epistemic perspective, while (b) the adoption of the narrator's. Hence, according to Domaneschi and Vignolo, only those participants selecting (b) express a truly descriptivist TVJ. Out of the total sample, 42% of the participants express the TVJ "True". However, 74% of them then select (a). Domaneschi and Vignolo infer that most of the supposedly descriptivist evidence does not provide genuine support for DT.

2.2.1.2 Experiment 2

As mentioned above, Domaneschi and Vignolo's study encompasses also two other conditions, differing in the final prompt. Experiments 2 and 3 consist of just one TVJ question, without any subsequent follow-up. In Experiment 2, participants receive the following prompt.

> That night, their dad asked: Do you know who won the Super Dog Race?
> Tom said: "Blaze was the dog that won the Super Dog Race".
> Emily said: "Pickles was the dog that won the Super Dog Race".
> Who tells the truth about the real (unbeknown to Tom and Emily's teacher) winner of the Super Dog Race?
> (a) Tom
> (b) Emily
> (c) Neither
> (Domaneschi and Vignolo, 2020, pp. 449-50)

Domaneschi and Vignolo design this final prompt to capture the narrator's perspective. The question asks participants who tells the truth about the «real» winner, who triumphed «unbeknown to Tom and Emily's teacher». Those words stress the extra information that the narrator provides and participants should rely upon when expressing their TVJ. If the *real* winner is *unbeknown* to Tom and Emily's teacher, then he/she will be unbeknown also to Tom and Emily themselves, as the teacher is their source of information about the race. Hence, the participant cannot respond to the final prompt based on Tom's or Emily's perspective, and must instead express her TVJ by adopting the information inaccessible to Tom, Emily and their teacher – namely, the supplementary information that the narrator provides. Tom's utterance serves as a control-check, as his sentence is false according to both DT and CHT. The different predictions of the two theories pertain to (b) and (c). According to DT, participants should choose (b). Emily tells the truth about the real winner of the race, as she associates the name "Pickles" with "The winner of the Super Dog Race", which picks out Max. Therefore, Emily expresses a true proposition about that dog. Instead, according to CHT, participants should choose (c), i.e. "Neither". By using the name "Blaze", Tom talks about Blaze. By using the name "Pickles", Emily talks about Pickles. Hence, neither Tom nor Emily expresses a true proposition about the real winner of the Super Dog Race, Max. Notably, 74% of participants choose (c), while 18% (b) and 8% (a). The substantial majority of participants opt for the purported causal-historical answer.

2.2.1.3 Experiment 3

Experiment 3 complements Experiment 2 by shifting the emphasis to the speaker's perspective. The final prompt is the following:

> That night, their dad asked: Do you know who won the Super Dog Race?
> Tom said: "Blaze was the dog that won the Super Dog Race".
> Emily said: "Pickles was the dog that won the Super Dog Race".
> Who reports correctly the story of the Super Dog Race that the teacher told to Tom and Emily?
> (a) Tom
> (b) Emily
> (c) Neither
> (Domaneschi and Vignolo, 2020, pp. 450)

The question asks whose utterance reflects the accepted consensus within the linguistic community to which Tom, Emily and their teacher belong, focusing on the epistemic viewpoint accessible to them. Thus, the participant should refrain from adopting the narrator's perspective, which encompasses information unavailable to the above community. Domaneschi and Vignolo predict that participants in this condition should mostly opt for (b), given that the teacher told at school: «Pickles was the dog that won the Super Dog Race». As expected, the vast majority of participants, 89%, select (b), while 5% (a) and 6% (c).

2.2.1.4 Conclusions

Domaneschi and Vignolo take Experiment 1, 2 and 3 to jointly support CHT. Experiment 1 reveals that the majority of the TVJs "True" derive from participants adopting the speaker's perspective. Hence, the TVJs "True" do not generally provide genuine support for DT. In contrast, the TVJs "False" provide genuine causal-historical evidence: given that a participant taking Emily's perspective will answer that her sentence is true, the TVJ "False" can only stem from a participant assuming the narrator's perspective and expressing a Kripkean answer. Therefore, Experiment 1 shows that, while a majority of the participants opting for TVJ "True" offer responses unsuitable for comparing DT and CHT, the remaining participants – those choosing the TVJ "False" – provide genuine

causal-historical evidence. In other words, all the data that is adequate to compare the two theories of reference supports CHT. Experiment 2 is designed to elicit TVJs from the narrator's perspective, the relevant one for testing theories of reference. The majority (74%) of these TVJs align with the causal-historical prediction that Emily's sentence is false. Experiment 3 elicits TVJs from Emily's epistemic perspective. The vast majority of participants, 89%, opt for the TVJ that in Experiment 1 aligns with DT, thereby further corroborating that the seemingly descriptivist evidence in Experiment 1 originates from participants adopting Emily's viewpoint. To echo Domaneschi and Vignolo's words:

> The results of Experiments 2 and 3 provide evidence that when the question task was disambiguated, participants gave responses consonant with Kripke's antidescriptivism. When the ambiguity of "true" as meaning *true from the narrator's epistemic perspective* or *true from Emily's epistemic* perspective was clarified, participants tended to understand Emily to refer to the referent predicted by antidescriptivism and not to the referent predicted by classical descriptivism.
> (Domaneschi and Vignolo, 2020, p. 452, emphasis in the original)

Domaneschi and Vignolo situate their conclusion within the context of their group of Italian participants. Nevertheless, their results suggest that the epistemic ambiguity might also affect Chinese participants and individuals from other countries. Based on their findings, «it would be naïve to rely on truth-value judgments of participants in other linguistic communities [like East Asians ones], regardless of having previously excluded the impact of the epistemic ambiguity» (p. 452).

2.2.2 A critical analysis

2.2.2.1 Attributive uses

A very appreciable feature of the Super Dog Race, which neither Li et al. (2018) nor Vignolo and Domaneschi (2022) emphasize, has to do with the sentence whose truth value participants have to judge. The sentence is "Pickles was the dog that won the Super Dog Race". The predicate of this sentence contains the description that, according to DT, fixes the name's referent – namely, "The dog that won the Super Dog Race". This feature precludes a possible attributive interpretation of the name. As argued (section 1.3.2), in the case of attributive uses, a causal-historical user suspends the Kripkean reference-

fixing mechanism of the name to refer to whoever satisfies the authorship associated with the name and to attribute some property related to that authorship. For example, a speaker may suspend the conventional reference-fixing mechanism of "Gödel" to refer to the author of the incompleteness theorem, whoever he/she is, and attribute the property of being a mathematical genius to that individual. One can understand attributive uses as a pragmatic phenomenon, in line with Kripke (1980), or as a semantic one, as Devitt (1981, 2015) suggests: in the former case, an attributive use is a non-literal phenomenon; in the latter, it corresponds to a secondary, yet literal meaning that a name acquires when it becomes firmly associated with a widely shared description. However, as Heck (2018, p. 258) stresses, taking a stand on this matter is beyond the point. Whichever explanation one adopts, attributive uses of names are a descriptive phenomenon that the causal-historical theorist *already* acknowledges. For example, Martí (2015) writes:

> In fact, it [the circumstance that speakers sometimes use proper names descriptively] is something that has been pointed out by one of the staunchest defenders of a causal theory of reference. Michael Devitt […] has often pointed out that proper names are used descriptively in some contexts. Literary historians, for instance, use "Shakespeare", bypassing anti-Stratfordian concerns, to talk about whoever wrote the famous plays, when their focus is the discussion of the plays and their contents.
> (Martí, 2015, p. 168)

To use Devitt's (2015) words:

> […] the names of authors […] can have a double life. In claims about where "Shakespeare" lived, was educated, and so on, the name seems to function as a designational [causal-historical] name. In critical assessments of "the works of Shakespeare", however, it often seems to function as a descriptive name, so that it would not matter to the truth of the assessments if the work was actually written by Bacon.
> (Devitt, 2015, p. 127)

In light of these observations, if a participant provides a descriptivist answer in linguistic contexts that allow for an attributive understanding of the name, that response does not constitute a threat to the Kripkean theory of names. However, as mentioned, the

final sentence in Li's Super Dog Race seems to preclude that possibility. A causal-historical user of names will understand a name as attributive in those contexts where an attributive use engenders new expressive possibilities. For example, a Kripkean speaker can use the name "Shakespeare" attributively to bypass the anti-Stratfordian concerns and ascribe some attribute, for example the property of having revolutionized the literature, to the individual who wrote certain works, whoever he/she is. However, the situation seems different in the context of a sentence like "Pickles is the dog that won the Super Dog Race". In this case, there seems to be no gain for Emily in using the name "Pickles" attributively. Under an attributive construal of the name, Emily would be suspending the name's conventional reference-fixing mechanism to talk about the dog – whoever he/she is – that won the Super Dog Race and then ascribe the property of being the winner of the Super Dog Race to that dog. Such an approach does not yield any expressive advantage. As argued, the crucial gain of an attributive use lies in the possibility of linking a certain authorship[67] to a property that is intrinsically related to it. For example, a speaker can connect the authorship of the incompleteness theorem to the property of being a mathematical genius because the individual who discovered the theorem must have been a mathematical genius. However, if the property that one ascribes is the same for which one suspends the conventional reference-fixing mechanism of the name, no expressive advantage emerges. In such a context, it is unclear why a causal-historical participant should understand the name attributively. Therefore, if a participant chooses the TVJ "True" (epistemic ambiguity aside), then she expresses an answer that the causal-historical theorist is not already disposed to accept. Hence, TVJ experiments in which the critical sentence's predicate encompasses the purportedly reference-fixing description have the advantage of obviating an attributive understanding of the name. Thus, such studies – other potential confounding factors aside – can gather genuine descriptivist answers.

[67] As said (section 1.3.2, n. 22), in this dissertation the notion of "authorship" takes a broad interpretation in connection to attributive uses: a person is the author not only of a literary, scientific or artistic work, but also of any other kind of achievement, including, for example, winning the Super Dog Race. For example, in a discussion about the Super Dog Race, a speaker may use "Pickles" attributively in the sentence "Pickles must have been well trained to run so much faster than the others". I thank Genoveva Martí for bringing out this point and providing the above example.

2.2.2.2 Perspectives, factual truth and non-literal interpretations of the truth predicate

As seen, Domaneschi and Vignolo (2020) maintain that the predicate "true" is ambiguous between the interpretation "true according to the speaker's perspective" and "true according to the narrator's epistemic perspective", as also the title of their article signals: "Reference and the ambiguity of truth-value judgments". However, Devitt and Porot (2018) and Devitt (2023) object to that thesis. When a speaker S asks whether a proposition P is true, S is not asking about *anyone's perspective* on the truth of P: S is inquiring about the *factual truth* of P. As Devitt and Porot (2018) stress, "if we ask whether it rained at Trump's inaugural, we are not asking whether Trump, or anyone else, *thinks* it rained" (p. 1555); rather, we are asking about a fact, that is, whether it rained at Trump's inaugural. Analogously, in the Super Dog Race scenario, the final TVJ prompt pertains to which dog factually won the race and not about anyone's opinion – i.e. perspective – on the matter. Therefore, the truth predicate is not ambiguous and «if a subject takes the question to be about anyone's perspective [e.g., Emily's] on the race, she has *mis*understood the question» (Devitt, 2023, p. 1152, emphasis in the original).

I overall agree with Devitt and Porot (2018) and Devitt (2023) – but I will specify some provisos as to the verb "misunderstand" at the end of this section. To use their example, if I ask John whether it rained at Trump's inaugural and he answers that "Smith thinks that it did", John is not directly answering my question. Maybe John does not know whether it rained or not and he is providing me with his most informed answer, reporting the opinion – i.e., the perspective – of a person whom he takes to be reliable. Therefore, John's question may be conversationally and pragmatically acceptable. However, that answer does not address my specific question, which pertains to the meteorological conditions during Trump's inaugural ceremony, rather than Smith's or anyone else's opinion on the matter.[68] Therefore, if a participant understands the TVJ final prompt from Emily's perspective, namely as asking whether Emily's linguistic community takes her sentence[69] to be true, then the participant is not understanding the question literally. Vignolo and Domaneschi (2022), following Sytsma and Livengood (2011), provide experimental data corroborating the thesis that, in the context of RIs, participants happen

[68] The metaphysical background for these considerations is a realist one, which Devitt himself (1997) supports, according to which there is a distinction between the way things actually are and the way people think things are.

[69] A terminological clarification is in order. According to prevailing interpretation, propositions are the primary bearers of truth and falsity (McGrath and Frank, 2020). However, for the mere sake of simplicity, in this dissertation I attribute truth and falsity also to sentences and utterances.

to take the phrase "the person who discovered the theorem" as equivalent to "the person who John thinks discovered the theorem". Domaneschi and Vignolo (2020) argue that a comparable phenomenon takes place in the context of TVJs as well.

As regards the participants adopting the so-called "narrator's perspective", Domaneschi and Vignolo take them as judging the factual truth of Emily's utterance.[70] However, one may question whether the label "narrator's perspective" is appropriate in this context. As discussed, the factual truth corresponds to the way things actually are, irrespective of anyone's opinion – i.e., perspective – about them. Domaneschi and Vignolo stress that the "narrator's perspective" is necessary to evaluate the sentence's factual truth. The narrator is the person providing a complete narrative of the Super Dog Race, which includes information to which Emily's community is not privy: first, the fact that the announcer nor anyone else got track of a dog who crossed the line first and run to the North Pole; second, the consequent fact that the description that Emily associates with the name does not pick out the original bearer of it. Those pieces of information are instrumental to assess the factual truth value of Emily's utterance. The descriptivist participant will be sensitive to the satisfactory relation between Max and the description that Emily associates with "Pickles". On the other hand, the causal-historical subject will prioritize the causal-historical path connecting the dog who crossed the line in second place and Emily's use of "Pickles". Therefore, in Domaneschi and Vignolo's construal, the label "narrator' perspective" does not imply the judgment of something different from the sentence's factual truth value. Instead, it emphasizes that the narrator is the source of a comprehensive and reliable account of the events, which is essential for assessing the factual truth value of Emily's statement.

Domaneschi and Vignolo are right that a participant judging the factual truth value of Emily's sentence needs to consider the entirety of the information that the narrator provides. However, I also share Devitt and Porot's concern that the label "narrator's perspective" might be somewhat misleading. The contrast is between the participant adopting the speaker's perspective on the one hand and the participant judging the factual truth on the other. The participant adopting Emily's perspective evaluates whether Emily's sentence aligns with the information to which Emily's linguistic community has

[70] To be precise, Domaneschi and Vignolo (2020) never explicitly state that the participant adopting the narrator's perspective judges the factual truth value of the sentence. However, that seems to be the idea underpinning their paper, as they confirmed in personal communication.

access. The participant judging the factual truth value of Emily's utterance evaluates not whether the sentence aligns with the so-called "narrator's perspective", but rather whether it aligns with the facts themselves. As Domaneschi and Vignolo (2020) correctly point out, a participant can express such a judgment only by relying upon the entire account that the narrator provides. Yet, one matter is the information that the participant uses to express her TVJ. Another matter is the nature of the TVJ itself: the TVJ concerns the facts themselves, and therefore, if the participant chooses the TVJ "True", she means that the sentence is factually true (i.e., true strictly speaking) and not true from the narrator's or anyone's perspective. For these reasons, the expression "factually true" is arguably more appropriate and less misleading than "true from the narrator's perspective".[71]

As mentioned, it is possible to harbor some reservations concerning the verb "misunderstand" in relation to those participants who adopt the speaker's perspective. The term "misunderstand" entails or at least suggests some unintentional error on the participant's part. It is possible that some participants misinterpret the truth predicate without realizing their mistake. However, it seems unduly critical to claim that a majority of Chinese – and a non-negligible minority of Westerners – run into that kind of mistake and thus are «incompetent survey takers» (Li, 2023b, p. 1160).[72] Rather, a plausible interpretation is that Chinese are *more charitable* survey takers: they are sensitive toward Emily's peculiar and limited epistemic condition that derives from her perspective on the events. Therefore, Chinese participants may lean toward deviating from the literal interpretation of the truth predicate to avoid judging Emily's utterance as false. In doing so, they adopt a non-literal reading of the truth predicate: *in that sense*, Chinese participants are more inclined to adopt a "misreading" of the truth predicate. However, they do so not because they are bad survey takers, that is out of a failure to comprehend the truth predicate, but rather because they are more sensitive to the peculiar and limited epistemic conditions in which Emily finds herself. Thus, Chinese participants try to find a way out to charitably judge her sentence as "true". Hence, an arguably more precise

[71] Devitt himself acknowledges the possibility that participants erroneously perform the experiment not by judging the factual truth but by adopting the speaker's perspective. Indeed, in an experiment on natural-kind terms, Devitt and Porter (2021, p. 30, n. 149) try to control for that factor. Therefore, Devitt is not against the epistemic-perspective talk *per se*. However, given the ontological status of facts, which are independent of any opinion or perspective on them, Devitt disagrees with the use of the label "narrator's perspective" to denote the approach of the participants judging the factual truth of the sentence.

[72] Note, however, that Li uses (also) that consideration to argue for a referentially pluralist position as regards the reference of proper names, which I do not endorse (section 2.3.2.3).

depiction of the phenomenon is not that some participants *misunderstand* the truth predicate, but rather that some participants *consciously adopt a non-literarily reading of it out of charity*. While the former formulation suggests some unintentional error on the participant's part, the latter states a deliberate departure from the conventional understanding of the predicate itself.

In light of all these reasons, I disagree with Domaneschi Vignolo's thesis that the truth predicate is ambiguous: if one asks someone whether the proposition P that a speaker S expresses is true, what one literally asks is uniquely whether P is factually true. There is no alternative meaning according to which the question asks whether S or S's linguistic community takes P to be true.[73] Nevertheless, Domaneschi and Vignolo's concern is legitimate: some participants may use the truth predicate by targeting something different from the factual truth, and their study provides at least some tentative evidence in that direction. Therefore, when in the forthcoming sections I talk of "epistemic ambiguity", this phrase will simply denote the two possible approaches based on which a participant can perform her TVJ task. Moreover, for the sake of simplicity, I will also refer to two possible "interpretations" of the TVJ question. However, I will take for granted that only the participants reading the question as one about the factual truth interpret it literally, while the other participants adopt a non-literal interpretation of the truth predicate.

Two final considerations are in order. First, the non-literal interpretation of the truth predicate seems to be part of a broader phenomenon that Holton (1997) identifies as "protagonist projection", in which people actively «project[…] [themselves] in the point of view of the protagonist» (p. 626). Building upon some insights that Holton elaborates, Buckwalter (2014) provides empirical evidence that the phenomenon of protagonist projection enables theorists to account for certain non-factive instances of the verb "to know" without abandoning the traditionally assumed factivity of the verb. For example, Buckwalter considers the following sentence: "Everyone knew that stress caused ulcers, before two Australian doctors in the early 80s proved that ulcers are actually caused by bacterial infection" (Buckwalter, 2014, p. 393)[74]. The statement seems perfectly

[73] I set aside peculiar cases, such as statements about taste, with regard to which the truth predicate may behave differently, because those cases do not pertain to the TVJ literature on the Gödel Case.

[74] Buckwalter draws the sentence from Hazlett (2010), who in turns adapts it from Achenbach (2005), who writes: «then in the early 1980s two Australian doctors figured out that peptic ulcers were caused not by stress but most often by bacterial infection. That was a stunning development. Everyone knew that stress caused ulcers! People would say: "This job is giving me an ulcer." No, the microbe did it».

acceptable. As the complement clause after "knew that" expresses a false proposition, this example seems to challenge the supposed factivity of the verb "to know". However, Buckwalter empirically defends a traditional, factive account of the knowledge predicate. He shows that, when asked to say whether – for example – people really knew that stress caused ulcers or rather they only thought they knew, participants opt for the latter explanation. Therefore, the participant refrains from literally ascribing knowledge in the case of false belief, but rather engage in the above-mentioned phenomenon of protagonist projection: the judger takes the perspective of the relevant protagonists – in this case, the people before the early 80s. Buckwalter writes:

> […] people think Know Ulcers [i.e., the sentence in question] is acceptable only because they assume the protagonist's perspective, in this case the perspective of putative knowers before [the discovery that stress does not cause ulcer]. Then, they describe how the situation may have appeared from that perspective. People living before the 1980s had relatively decent evidence supporting their beliefs about the relationship between stress and ulcers, in a way typically thought to be consistent with knowledge. Therefore 'everyone knew that stress caused ulcers' is judged acceptable because it appeared acceptable from that perspective – and not because it's judged true, from the more informed evaluator's perspective, that the protagonists actually knew something false.
> (Buckwalter, 2014, p. 395-6)

A conceptually analogous phenomenon characterizes the approach of those participants who perform the TVJ final prompt of the Gödel Case from the speaker's perspective. As in the above ulcer-related scenario the participants «describe how the situation may have appeared from that [the protagonists'] perspective», similarly in the Gödel scenario participants judge Emily's utterance by evaluating the truth value that Emily and her community ascribe to it.[75]

The second final consideration is this. Even if one, *pace* what I have argued so far, adopts a relativistic stance and contends that the truth predicate is literally ambiguous

[75] See also Domaneschi and Di Paola (2019), who provide further evidence corroborating the thesis that the phenomenon of protagonist projection is what guides the non-factive uses of the verb "to know". Moreover, they argue that the Relevance Theory (Sperber and Wilson, 1995; Carston, 2002) offers the most compelling framework to explain the psychological mechanisms driving the interpretation of non-factive knowledge attributions.

between “true according to someone’s perspective” and “factually true”, the problem that perspectivism raises persists. One of the two meanings (i.e., “true from the speaker’s perspective”) is simply irrelevant to making empirical tests, as *any* theory of reference agrees that – when understood in that way – Emily’s utterance is “true”. Two theories of reference can instead conflict as to the factual truth of the proposition that Emily expresses. Therefore, any experimental test must to ensure that participants approach the TVJ final prompt by adopting the factual-truth interpretation.[76]

2.2.2.3 Experiment 1

One may object that both the follow-up options in Experiment 1 are inaccessible to the descriptivist participant. I will argue that this is a problem only in the context of the follow-up option whose goal is to capture the answers of the participants judging the factual truth value of Emily’s sentence. However, let us first consider the follow-up option whose aim is to identify the answers of the participants adopting Emily’s perspective:

> Pickles did not win the Super Dog Race, but Emily believes that Pickles won the race because her teacher told it to her.
> (Domaneschi and Vignolo, 200, p. 449)

This option consists of two parts. The first states “Pickles did not win the Super Dog Race”, while the second asserts “but Emily believes that Pickles won the race because her teacher told it to her”. The occurrence of “Pickles” in the first part of the sentence refers to Pickles, as that is the dog that the announcer mistakenly presents as the winner of the race. This thesis gains reinforcement from the narrator’s consistent usage of the name “Pickles” throughout the vignette to refer to that specific dog. Also the second occurrence of “Pickles”, appearing in the subsequent part of the follow-up option, likely refers to the same dog, as a sentence using the same name with a changing referent would be deviant. The potential problem is that, according to DT, Emily *cannot* hold a belief

[76] Against the relativistic interpretation of the truth predicate, see the experimental work of Barnard and Ulatowski (2021), which indicates that laypeople’s common understanding of truth is objectivist: «attributions of relativism about truth to ordinary people seem out of place, given what we’ve found in this experiment» (p. S730). See also Ulatowski (2018).

about that dog. It is beyond doubt that Emily holds the verbalized belief that "Pickles won the Super Dog Race". However, that belief can have Pickles as its object only if the occurrence of "Pickles" appearing therein gets its referent through the name's causal-historical chain. In that case, yet, the name's chain of communication would acquire semantic significance, which is exactly what DT denies. According to DT, Emily's belief is about Max, as the name "Pickles" appearing therein gets its referent through the description that Emily associates with it, that is "The winner of the Super Dog Race".[77] A descriptivist participant can take this option to be accessible only if she interprets its wording as stressing Emily's *de dicto* representation, namely, the terms that Emily would use to express her belief. Undoubtedly, Emily would utter: "Pickles is the winner of the Super Dog Race". The adoption of this *de dicto* interpretation is not implausible, but it requires the descriptivist participant, within the context of the follow-up answer, to understand the first occurrence of "Pickles" as referring to Pickles, while the second as referring to Max. A more straightforwardly accessible formulation for the descriptivist would have been "Emily reports what the teacher told her". This wording uses the terms appearing in the final prompt of Experiment 3, which Domaneschi and Vignolo specifically design to stress Emily's perspective.

However, as mentioned, the inaccessibility of this follow-up answer to the descriptivist arguably does not constitute a problem. For the descriptivist participant, Emily's sentence is not only one that Emily and her community take to be true but is also factually true. Given that the sentence is factually true, the descriptivist participant has no reason to adopt a non-conventional (mis)interpretation of the truth predicate and restrict herself to the claim that Emily's sentence is true in relation to what a certain linguistic community believes. Therefore, the descriptivist participant should opt for the alternative follow-up option, which stresses that Emily asserts a factually true sentence because she is talking about the dog that actually won the race. On the contrary, the causal-historical participant may have reason to adopt the unconventional interpretation of the final prompt, as it enables her to charitably judge a sentence that is factually false as "true". In light of these reasons, the fact that the option stressing Emily's perspective is accessible only to a

[77] The seminal paper about the externalist individuation of mental contents, like beliefs, is Burge (1979). For a review of the literature on the topic, see Rowlands (2020).

causal-historical participant correctly reflects the expectations that a researcher can legitimately have in relation to the commitments pertaining to DT and CHT.[78]

More problematic, however, is the circumstance that the alternative follow-up option, the one that justifies the TVJ "True" with purportedly genuine descriptivist terms, is not straightforwardly accessible to the descriptivist participant. That option states:

> What Emily believes is a true description of Max and she uses the name "Pickles" to talk about the dog that really (unbeknown to Tom and Emily's teacher) won the Super Dog Race.

Domaneschi and Vignolo introduce the parenthetical remark "unbeknown to Tom and Emily's teacher" to emphasize the participant's privileged information as opposed to the limited epistemic status of Emily's linguistic community. This option should capture the descriptivist participants who express the TVJ "True" for genuinely descriptivist reasons. However, if DT is the correct theory of reference, Max is not the dog that won the Super Dog Race "unbeknown to Tom and Emily's teacher". Emily and Tom's teacher does know that Max won the Super Dog Race. The teacher has the belief that "Pickles is the dog that won the Super Dog Race". As argued, if DT is the correct theory of reference, the occurrence of "Pickles" in that belief refers to the dog that won the Super Dog Race, namely Max. Hence, the teacher's belief is *about* Max and is true, as Max is the winner of the race. Moreover, the belief is arguably justified, since the teacher formed it by reading the relevant newspapers and not – for instance – by simply guessing. Therefore, the teacher holds a true and justified belief, and thus possesses knowledge about Max and his victory in the Super Dog Race.[79] Therefore, this follow-up alternative is biased against DT, and a descriptivist participant may have troubles accessing it.

---

[78] I thank Massimiliano Vignolo for his precious insights on this point.

[79] The seminal work exploring the necessary and sufficient conditions defining knowledge is Gettier (1963), where two famous thought experiments would show that a justified true belief alone is insufficient for constituting knowledge, and that some further condition is required. Therefore, one may object that arguing that the teacher's belief is true and justified does not prove that it amounts to knowledge. However, the acceptance of true justified beliefs as sufficient for knowledge formation is not essential for my argumentative purposes. The central focus of my criticism is that, under the assumption that DT is true, Domaneschi and Vignolo's remark that Max's victory is *unbeknown* to Tom and Emily's teacher is too strong and hurried, and therefore not easy to process for a descriptivist participant. Gettier's article has sparked a substantial body of critical literature: for a comprehensive survey of the subject matter, see Ichikawa and Steup (2018). Important elaborations on the Gettier scenarios also emerged in experimental philosophy, commencing with the pioneering study of Weinberg et al. (2001); for further developments, see Machery et al. (2015a), Machery et al. (2017), Rose et al. (2017); for reviews of the literature, see Turri (2015, 2018) and Beebe (2024).

A descriptivist language user can make the alternative consistent with her understanding of names by simply ignoring the parenthetical remark "unbeknown to Tom and Emily's teacher". In that case, the option becomes straightforwardly intelligible. Alternatively, a descriptivist can interpret that parenthetical remark as signaling that the *modality* in which Max won the race has remained undisclosed to the teacher. The teacher does not know that, after crossing the line first, Max felt an overwhelming sense of pride, continued running until he reached the North Pole and subsequently disappeared without anyone seeing him again. Such a reinterpretation of the parenthetical remark or the omission of it does not require a profound restructuring of the sentence, and thus a descriptivist participant can arguably engage with it. That said, the necessity of these alternative reinterpretations to make the option consistent with DT indicates that the wording is less than optimal and that the final prompt is biased against the descriptivist view.

More in general, the problem that I identify with the follow-up options of Experiment 1 reflects some flaws in the formulation of the vignette itself that Domaneschi and Vignolo draw from Li et al. (2018). The narrator uses the name "Pickles" to talk about the dog that the announcer mistakenly took as the winner of the race, since the vignette states that "Pickles crossed the finish line second". As "Pickles" in the narrator's language is a name of the non-winning dog, the narrator – if DT is the correct theory – should not be willing to utter that "Tom and Emily learned at school that Pickles won the Super Dog Race". When hearing the name "Pickles", the people within Emily and Tom's linguistic community do not acquire any information about Pickles, as the occurrences of "Pickles" within their community refer to Max. Therefore, Max is the dog about whom Tom and Emily learn something when they listen to their teacher's lesson. For consequential and specular reasons, the narrator should not assert that Tom and Emily "don't know anything about Max", because, as discussed, their beliefs are about Max and they likely amount to knowledge.[80] Also in this case, a possible way in which the descriptivist participant can

[80] While in section 2.1.1.3 I have argued that MMNS's supposition that "Gödel did not discover the incompleteness theorem" is admissible under the adoption of Reference Descriptivism, it is worth noting that other points of MMNS's vignette, similar to Li et al.'s (2018) text, are incompatible with any version of descriptivism – causal descriptivism aside. For the reasons presented in the main text, a descriptivist narrator should not be disposed to utter that "the claim that Gödel discovered the incompleteness theorem is the only thing they [John and his fellows] have ever heard *about Gödel*» (p. B6, my emphasis). Also the statement «His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed *to Gödel*» may be subject to dispute (*ibidem*, my emphasis). According to DT, it is not true that John and his fellows attribute the work to Gödel (i.e., the thief). Were they asked "To whom do you attribute the incompleteness theorem?", they would reply "Gödel", which in their language picks

take the text as consistent with DT is to assume that those belief attributions – in terms of what Tom and Emily "learned" or "know" – convey the *de dicto* representations that those characters hold. Put differently, the vignettes present the words with which Tom and Emily would express what they learned or what they know. It is beyond doubt that, if someone asked Emily or Tom to say what they learned, they would use the name "Pickles" and utter that "Pickles won the Super Dog Race". Analogously, should they be asked about whom their knowledge is, they would reply "Pickles".[81]

out Schmidt, as their reference-fixing description is "The discoverer of the incompleteness theorem". However, additional factors make the situation slightly more intricate. During Gödel's lifetime, the theorem was in fact attributed to him. At that time, speakers plausibly associated "Gödel" with information extending beyond the sole discovery of the incompleteness theorem, thereby delineating an overall reference-fixing description of which Gödel was likely the best satisfier. Therefore, in the language of those speakers in the past, "Gödel" was a name of the thief, and therefore he was the individual to whom they in fact attributed the discovery of the theorem, *although* he is no longer the individual to whom the vignette's contemporary speakers (i.e., John and his fellows) attribute the result. Similar considerations extend to the announcer in Li et al.'s (2018) vignette, who in the past declared Pickles the winner of the race: unlike the contemporary speakers, the announcer and his community plausibly associated "Pickles" with some description(s) picking out the dog that in fact crossed the line second. I am grateful to Michael Devitt for precious reflections on this topic.

[81] Ding and Liu (2022) commit a conceptually analogous mistake when they argue that, if the reference-fixing description that John associates with "Gödel" were "The person who is *widely believed* to have discovered the incompleteness theorem", instead of "The person who discovered the incompleteness theorem", DT could account for the intuition that John is talking about the thief. However, that is not so (D'Agruma, manuscript). If the reference-fixing description that John associates with the name were "The person who is widely believed to have discovered the theorem", the descriptivist participant should figure out the referent of "Gödel" based on how John and his fellows would respond to the question "Who do you believe discovered the incompleteness theorem?". They would answer "Gödel", thus showing that they have the belief that "Gödel is the discoverer of the incompleteness theorem". At this point, the participant needs to determine the referent of "Gödel" appearing in this belief. As a descriptivist, she will resort again to the name's reference-fixing description, that is "The person who is widely believed to have discovered the incompleteness theorem", engendering a circularity. Hence, under the assumption that DT is the correct theory of reference and that the aforementioned description fixes the reference of "Gödel", the name has no referent.

2.2.2.4 Experiment 2

Concerns similar to the ones that I have just presented extend to Experiment 2. Let us recapitulate the final prompt for easy reference:

> That night, their dad asked: Do you know who won the Super Dog Race?
> Tom said: "Blaze was the dog that won the Super Dog Race".
> Emily said: "Pickles was the dog that won the Super Dog Race".
> Who tells the truth about the real (unbeknown to Tom and Emily's teacher) winner of the Super Dog Race?
> (a) Tom
> (b) Emily
> (c) Neither
> (Domaneschi and Vignolo, 2020, pp. 449-50)

This final prompt encounters the same problem afflicting the follow-up option that in Experiment 1 stresses the so-called "narrator's perspective", namely, the parenthetical remark "unbeknown to Tom and Emily's teacher". As discussed, according to DT, the teacher as well as Tom and Emily are aware of Max, namely the real winner of the Super Dog Race, and they know that he won the race. Thus, some potential descriptivist participants may have answered "Neither" simply because they found the question itself difficult to comprehend, approximating self-contradictoriness in terms. On the one hand, a descriptivist participant should recognize that Tom, Emily and their teacher know of the real winner of the race. However, on the other hand, the final prompt asserts that the winner is unbeknownst to Emily and Tom's teacher. Once again, a way out for the descriptivist is to simply ignore the parenthetical remark or interpret it as regarding the modalities with which Max won the race. Such reinterpretation of the sentence is not profound, and therefore it is a way out that a descriptivist participant can plausibly adopt. That said, the fact that a descriptivist participant has to resort to it shows that the formulation is biased against DT and thus less than optimal.[82]

[82] Machery (2021) too criticizes Domaneschi and Vignolo's (2020) Experiment 2. He writes: «The probe question of Experiment 2 is obviously leading participants to give answer C ["Neither"] since Max is the dog that Tom and Emily's teacher are not aware of» (p. 545). According to Machery, the problem inherent to Domaneschi and Vignolo's final prompt is that it asks about the dog Tom and Emily's teacher is not aware of, thereby leading participants to focus on Max. Instead, my criticism is that that dog won the Super Dog Race unbeknown to Tom and Emily's teacher only according to CHT, and thus the prompt encompasses a bias against DT. Li (2021) seems to raise a criticism similar to Machery's, although it is arduous to provide a more comprehensive assessment of it because Li simply restricts herself

2.2.2.5 The follow-up question is theoretical

The potential bias against DT is not the primary limitation of Domaneschi and Vignolo's (2020) study. Rather, my main concern about the experiment lies in the fundamental reliance upon the follow-up question of Experiment 1, which is arguably susceptible to the criticism that Martí (2009) raises against the use of referential intuitions. As discussed, Martí objects that MMNS (2004) is problematic because participants may express their RI not based on their linguistic usage, but rather on their theoretical inclinations: their replies would simply reflect the theoretical view – causal-historical or descriptivist – that they find intuitively more compelling. One may expand upon Martí's criticism and argue that participants in Domaneschi and Vignolo (2020) might answer the follow-up question by merely choosing the explanation that they find theoretically more intuitive. Let us recap the follow-up question and options:

> You think Emily's claim is true because:
> (a) Pickles did not win the Super Dog Race, but Emily believes that Pickles won the race because her teacher told it to her.
> (b) What Emily believes is a true description of Max and she uses the name "Pickles" to talk about the dog that really (unbeknown to Tom and Emily's teacher) won the Super Dog Race.
> (Domaneschi and Vignolo, 2020, p. 449)

The follow-up options provide two different theoretical frameworks to justify the TVJ "True": (a) centers on what Emily believes, while (b) relies upon the referential connection between the name "Pickles" and the real winner of the race. As in MMNS (2004) one may choose a certain referential intuition due to its intuitive nature even when it does not accurately reflect her own actual use of names, similarly in Domaneschi and Vignolo (2020) participants may choose – for instance – the theoretical explanation focusing on Emily's perspective because it appears intuitive, even when it does not accurately reflect the genuine reason behind their TVJ. In other words, a descriptivist participant may provide the TVJ "True" on genuinely descriptivist grounds; however, she

to the claim that «[…] in [Domaneschi and Vignolo's (2020)] experiment 2, the formulation of the test question biased participants towards the experimenter's expected answer» (p. 422).

may then select the follow-up answer (a) because she finds the explanation therein theoretically more intuitive and easier to process. After all, it is beyond doubt that Emily believes that Pickles won the Super Dog Race – under a *de dicto* interpretation of Emily's belief.[83] Therefore the participant may select that option over the more convoluted (b), which involves difficult expressions like "true description of" and "unbeknown to Tom and Emily's teacher", as well as the phrase "[Emily] uses the name", which precisely corresponds to the RI language which Martí warns researchers against. Machery (2021) expresses a somewhat analogous concern about Domaneschi and Vignolo's (2020) Experiment 1: «unfortunately, Experiment 1 suffers from severe demanding characteristic: asking participants Question 2 [i.e., the follow-up question] invites them to rethink their answer to Question 1. So, even if participants had descriptivist leanings when answering Question 1, they would be pushed to give a causal-historical answer when answering Question 2» (p. 545).[84]

In conclusion, it is possible to summarize the problem at hand as follows: scholars introduced TVJs as a means to tackle the inherent theoretical nature of RI tests. However, TVJs may be epistemically ambiguous. In an effort to address that problem, Domaneschi and Vignolo use a follow-up question. However, the follow-up question exhibits the same theoretical nature afflicting RIs, thereby vinifying the very reason scholars resorted to TVJs in the first place.

[83] One may object that a descriptivist participant cannot provide the TVJ "True" on descriptivist grounds and then take (a) as providing an intuitively compelling theoretical framework to justify that TVJ, since – as seen – (a) presents a wording that is inaccessible to the descriptivist – unless the subject opts for a de *dicto reading* of it. However, as argued, also (b) is consistent with DT only under a reinterpretation of its wording. Therefore, on this front, both options are equally flawed.

[84] One may wonder whether the criticism that Machery (2021) raises against Domaneschi and Vignolo (2020) is consistent with his own view according to which referential intuitions are at least as reliable as linguistic-usage measures like TVJs (Machery, 2011b; Machery, 2012b; Machery, 2014; Machery, 2021; Machery and Stich, 2012; Machery, Olivola and de Blanc, 2009). As said, option (b) in Domaneschi and Vignolo's (2020) Experiment 1 basically consists in a descriptivist referential intuition. If a participant is a descriptivist and her referential intuitions are as reliable as her TVJs, then she should first choose the TVJ that DT predicts and then justify it via the descriptivist referential intuition. In other words, the two sources should converge as they are equally reliable evidence regarding the same phenomenon. Whether Machery's criticism against Domaneschi and Vignolo (2020) is consistent with his own view on referential intuitions depends on the exact contours that one ascribes to Machery's criticism. One may interpret the problem with Domaneschi and Vignolo's (2020) Experiment 1 not to stem from the fact, *per se*, that (b) encompasses a descriptivist referential intuition. Rather, the issue may lie in the circumstance that the experiment's design solicits the referential intuition *after* a TVJ, introducing the «severe demand characteristics» that Machery mentions. Alternatively, one may argue that the problem lies in the wording of the descriptivist referential intuition (option b), which goes beyond a simple assertion such as "Emily is talking about Max", and instead uses a more complex formulation.

2.2.2.6 Social interpretation of predicates

Domaneschi and Vignolo's experiment faces an additional concern. Domaneschi and Vignolo borrow a simplified version of a scenario that Li et al. (2018) use. The scenario recounts the story of a dog, Pickles, that the announcer mistakenly takes as the winner of the Super Dog Race. However, the real winner is another Dog, Max, who, in his exuberance, continued running beyond the finish line and eventually reached the North Pole. As a consequence, no one, including the race announcer, noticed him and his victory. Following Li et al., Domaneschi and Vignolo claim that "The winner of the Super Dog Race" picks out Max. However, that claim is disputable. A predicate like "winner" can have two interpretations: one focuses on the brute facts, the other on the social reality. Under the former construal, Pickles did not win Super Dog Race and Max is the winner: as a matter of fact, the dog crossing the line first is Max. However, under the latter interpretation, Pickles is the winner of the Super Dog Race: the race announcer declared that dog to be the winner and the announcer is the person who holds the social authority to bestow such accolades. Social interpretations of predicates frequently emerge in the realm of sports. For example, suppose that a football player scores a goal while being in an offside position, yet the referee fails to notice the offside violation. There is a sense, centered upon the brute facts, in which that is *not* a goal. If a supporter of the opposing team exclaims "That wasn't a goal at all!", his utterance is sensible and intelligible. Nevertheless, there also clearly is a sense in which that goal *is* a goal: the match score changed as a consequence, prompting both teams to adjust their strategies accordingly – for example, the team that conceded the goal may adopt a more offensive approach to score one themselves. In that case, the social interpretation of the predicate "Being a goal" becomes operative. According to that construal, whatever the referee – namely the person in the relevant social position – declares as a goal is a goal. In sport contexts, the social construal of terms and descriptions seems even the most crucial one. If the predicate "Being the winner of the Super Dog Race" can be subject to a social interpretation, then there is another reason, in addition the potential adoption of Emily's epistemic perspective, as to why the TVJ "True" may not actually express a genuine descriptivist commitment. Also a causal-historical participant can answer that Emily's sentence "Pickles is the dog that won the Super Dog Race" is true if she understands the predicate

through a social lens. Indeed, Pickles – i.e., the dog originally named "Pickles" – is the competitor whom the announcer declared the winner of the race.

Li et al. (2018) are aware of the problem: «actions like winning a race or discovering an artifact could be socially or externally defined, making them somewhat ambiguous» (p. 110). The plausibility of a social interpretation of the predicate gains additional support from the justificatory comments that some participants express at the end of Li's (2023a) Study 3, which is a TVJ study similar to Li et al. (2018), with American participants only.[85] These participants «remarked that winning a race needs to be announced by the authority and crossing the finish line first without being witnessed by anyone does not count as a win» (p. 109). It is worth noting that the Super Dog Race is not the only vignette in Li et al. (2018). Another scenario, The Discovery of Claymen, describes a story in which Don digs a hole into the ground in search of water, but unexpectedly unearths valuable clay figures. Don meticulously records his findings in a notebook. Due to an emergency, Don leaves the city forever, inadvertently leaving the notebook behind. Subsequently, Peter steals the notebook and arrives at the exact location where Don dug the hole. As a result, Peter gets the credit for the discovery of the clay men. The vignette concludes that anyone today associates "Peter" only with "The discoverer of the clay men" (the details of the story are from Li, 2023a, p. 51)[86]. Interpreting this predicate in social terms means to understand "The discoverer of the clay men" as referring to the «the person who first shared that knowledge [i.e., regarding the clay men] with others rather than the person who initially uncovered the clay figures» (Li et al., 2018, p. 110). In other words, the discoverer would be the person responsible for the inclusion of the statues in museums and the subsequent social appreciation of them, rather than the person who factually encountered them initially. The social interpretation of this predicate may not appear immediately evident, especially when compared to predicates pertaining to sport contexts like the Super Dog Race. However, Li (2023a, p. 109) notes that some participants' explanatory comments in her Study 3, which employs The Discovery of Claymen as well, emphasize a social understanding of the predicate.[87]

[85] I analyze Li's (2023a) Study 3 more in detail in section 2.3.2.5.

[86] In this and the forthcoming sections, I often mention Li (2023a) while discussing Li (2021) and Li et al. (2018). Li (2023a) examines those experiments by exploring some unaddressed potential criticisms and offering new valuable conceptual insights.

[87] One may wonder whether an analogously problematic social interpretation extends to the predicate "Being the discoverer of the incompleteness theorem". I do not think so. Unlike discovering a statue, discovering a difficult theorem like the incompleteness one is not just a matter of being the first to stumble upon it. It entails possessing the

In light of this problem, Li (2021) uses four scenarios whose central events are less easily susceptible to a social construal. I discuss those vignettes and the related results in the forthcoming section.

2.3 Li (2021)

2.3.1 The study

The study in Li (2021) consists of three conditions: Original, External and Internal, in a between-subjects design. The Original Condition comprises a classical TVJ task; the Internal Condition stresses the speaker's perspective; the External Condition stresses what Li calls the "narrator's epistemic perspective". Structurally speaking, Li (2021) bears similarities to the work of Domaneschi and Vignolo (2020). However, as I will show, Li's experimental framing of the External Condition diverges from the one of Domaneschi and Vignolo's Experiment 2, which also has the aim to emphasize the so-called "narrator's perspective".

2.3.1.1 Original Condition

In section 1.4.3.2, I have briefly presented the Original Condition. Let us now revisit it and the corresponding results more in detail. Li uses four different vignettes, employing English versions for American participants and Mandarin versions for Chinese participants. As a way of example, I reproduce The Hospital in Newrock below:

> Long ago, a very strong man named Basil lived in Newrock. Basil was shy, so nobody knew about his power. Newrock needed a new hospital, but there wasn't enough money to pay for it. Basil wanted to help, so one night he carried in some wood and built a hospital all by himself. He was so strong that it only took him one night to finish. The next morning, Basil moved out of Newrock without telling anyone about the hospital. He never came back.
>
> That morning, a builder named Flint found the new hospital and went inside to explore. Just then, the town's mayor, George, walked by and saw Flint inside. He thought that Flint built the hospital, so George told the whole town. Soon everyone

intellectual brilliance necessary to develop a proof that leads to such a remarkable outcome. Therefore, it is more difficult to understand "Being the discoverer of the incompleteness theorem" under a social construal. The person who first shared the theorem – i.e., the thief – lacks the cognitive acuity that people crucially associate with such incredible scientific results. Being the *discoverer* of the incompleteness theorem also means being the *author* of it. The same conceptual affinity does not apply to the discovery of the clay men, as discovering them differs from being their author.

> knew about Flint's accomplishment. Flint, who couldn't hear or speak, couldn't tell them otherwise. Today, people still learn that Flint built the hospital in Newrock, but this is the only thing they have ever heard about Flint. No one has ever heard of Basil.
> (Li, 2021, p. 426)

The other three vignettes share a similar structure but differ in the central events that they portray. Instead of focusing on the building of a hospital, those scenarios revolve around the activities of drawing pictures, baking cakes and writing books. Li selects those predicates because she considers them less susceptible to a social interpretation when compared to those related to winning a race or discovering clay men. Naturally, a participant may try to develop a social interpretation for the above predicates. For instance, just as a participant can consider the discoverer of a statue to be the one who first made its existence known to the community, similarly one can consider the "builder of a hospital" to be the individual who first informed the community about its presence, ensuring that anyone in need of medical assistance could access it. The crucial factor in this context is the plausibility of interpreting such predicates through a social lens. On an intuitive basis, I agree with Li's assessment that predicates like "building a hospital", "drawing pictures", "baking a cake" and "writing a book" are less socially interpretable when compared to alternatives like "discovering clay men" or, even to a greater extent, "winning a race".

The final prompt that the participants answer in the Original Condition is:

> Emily learned at school that Flint built the hospital in Newrock. Like everyone else, this is the only thing she has ever heard about Flint. She has never heard of Basil.
> That night, Emily's dad asked her who built the hospital in Newrock.
> Emily said, "Flint was the person who built the hospital."
> Question: Is Emily right?
> (Li, 2021, p. 426)

As elucidated in section 1.4.3.2, the assessment of whether an individual is right in her utterance is conceptually equivalent to determining the truth or falsity of the expressed proposition. Therefore, Li's final prompt corresponds to a TVJ task. After expressing their TVJs, participants have to provide a justificatory remark clarifying the rationale

behind their choice. The results confirm the cross-cultural variation: across the four vignettes, Americans choose the TVJ "False" 90% of the time, while Chinese only 32%. As to the follow-up remarks, those participants – American or Chinese – opting for the TVJ "True" justify their choice by stressing Emily's epistemic perspective and the community's consensus:

(1) "Emily is right because that is what she learned. The first person in these stories should stop leaving the town forever right after they do something."
(2) "Emily, as a kid, was easily influenced by her surroundings, she said what she was taught."
(3) "We cannot blame the innocent people. From Emily's perspective, she is right."
(Li, 2021, p. 430)[88]

Conversely, those participants choosing the TVJ "False" stress what Li calls the "narrator's epistemic perspective":

(4) "From an outsider's perspective, Emily is not right because Basil was the individual responsible for the hospital."
(5) "Emily is not right because we as readers know the real person who built the hospital, which was Basil."
(6) "Emily is not right because factually, Walter wrote the book. Just because she is misinformed does not change the facts."
(Li, 2021, p. 431)

While in Domaneschi and Vignolo (2020) the label "narrator's perspective" is a – misleading – way to talk about the factual truth, in Li (2021) the precise contours of such a perspective are blurrier. The forthcoming sections also show where the vagueness in her stance lies.

[88] Some of these comments do not encompass the names "Basil" and "Flint". This is due to the fact that, while I have used The Hospital in Newrock as the illustrative scenario, the remarks that Li provides also derive from the other cases, which involve different names and predicates (e.g., writing a book). However, the narrative structure of the stories remains consistently analogous throughout.

2.3.1.2 Internal Condition

The Internal Condition focuses on the epistemic perspective of the vignette's linguistic community, which includes Emily as a member. Presented below is the final prompt of the Internal Condition:

> Last week, Emily learned at school that Flint built the hospital in Newrock. Like everyone else, this is the only thing she has ever heard about Flint. She has never heard of Basil.
> This week, Emily's teacher asked the class who built the hospital in Newrock.
> Emily said, "Flint was the person who built the hospital."
> Question: Will the teacher give Emily extra credit?
> (Li, 2021, p. 426)

Asking whether Emily will receive extra credit amounts to asking whether she correctly reports what the teacher told her and, more in general, the received consensus within her linguistic community. Only in that circumstance will Emily receive extra credit. Therefore, the final prompt encourages participants to adopt Emily's epistemic viewpoint. Also in this condition, following their response, participants have to justify their TVJ by providing a concise open follow-up explanation. The majority of participants, comprising 67% of Americans and 77% of Chinese, answer that Emily will receive extra credit. Participants justify this answer with explanations that emphasize Emily's epistemic perspective, which relies upon what the teacher told her. The following illustrative remarks provide a glimpse into the range of responses across the various vignettes:

> (7) "She gave the commonly accepted correct answer to the question of who drew the pictures."
> (8) "Emily's answer agrees with the teacher's belief."
> (9) "I think Emily would receive extra credit because everyone in town thinks that Honey baked the cake."
> (Li, 2023a, p. 119)

A non-negligible minority in each group answer that Emily does not deserve extra credit, comprising approximately one-third within the American sample and one-quarter within the Chinese one. Those participants provide justificatory remarks to the effect that «Emily did not deserve extra credit for her giving a simple answer to the teacher and failing to do thorough investigation» (p. 431). It seems that «this latter subgroup of participants did not fully understand the requirement of the credit-giving judgment task or they simply did not accept the rationale of this award-winning task» (Li, 2023a, p. 119). In other words, by emphasizing the simplicity of Emily's answer, those participants seem to situate the teacher's question within a context where students are expected to provide not only the name of the person who achieved the result (e.g., building the hospital), but also to conduct deeper research aimed to disclose additional details about that person (e.g., that he left the city after the construction of the hospital). An alternative explanation is that those participants try to articulate a causal-historical remark, by stressing that Emily should have undertaken comprehensive research and discovered that Basil, as opposed to Flint, actually built the hospital.[89] In any case, the answers of those participants deviate from the specific objective that Li pursues, namely assessing the extent to which participants perceive Emily's answer as aligning with the prevailing consensus within her community.

2.3.1.3 External Condition

The External Condition of Li's study aims to place emphasis on what she calls the "narrator's perspective". As mentioned, Li's External Condition differs from Domaneschi and Vignolo's Experiment 2, which also seeks to highlight the same perspective. This

[89] In this context, it is worth emphasizing a significant difference between DT and CHT. According to DT, Emily cannot discover that "Flint is not the builder of the hospital", because, in the language of her community, "Flint" refers to whoever built the hospital. As indicated in the main text, according to DT, what Emily can discover through more detailed research is for example that the builder of the hospital left the city immediately after completing the construal. Of course, she can also come to learn that the mayor mistakenly took another person as the builder of the structure, but Emily could not articulate that discovery by stating that "Flint was mistakenly taken as the builder of the hospital". What she should assert is: "The man who at the time was known as 'Flint' did not build the hospital; the hospital was built by a man called at the time 'Basil', but currently called 'Flint'". The report may appear convoluted, but that is how – if DT is true – Emily should describe her discovery. On the contrary, according to CHT, Emily can indeed discover that "Flint is not the builder of the hospital", because "Flint" refers to the same person who in the past was known by that name. For reflections on similar points, see Pérez Otero (2017).

divergence reflects different interpretations of what the "narrator's perspective" captures. Li's final prompt is as follows:

> Emily learned at school that Flint built the hospital in Newrock. Like everyone else, this is the only thing she has ever heard about Flint. She has never heard of Basil. Kermit read the story about the hospital and learned all about Basil and Flint. He also learned about what Emily had learned about the hospital in school.
> One day, Kermit's dad asked him who built the hospital in Newrock.
> Kermit said, "Flint was the person who built the hospital."
> Question: Is Kermit right?
> (Li, 2021, p. 426)

The results show a clear prevalence of the TVJ "False": Americans opt for it 95% of the time, while Chinese 85%. The participants justify their TVJ by stressing that Kermit learned the story from the vignette, in which the narrator provides a complete overview of the events. Stated below are some justificatory remarks:

> (10) "Kermit is not right because Walter originally wrote the book."
> (11) "Kermit read the story about Basil and Flint, and he knew that Basil had actually been the one who built the hospital."
> (12) "Kermit knows the existence of Bob. He is aware that Bob drew these pictures so to say that Sage drew them is incorrect."
> (Li, 2023a, p. 120)

While in their Experiment 2 Domaneschi and Vignolo try to stress the so-called "narrator's perspective" by emphasizing the privileged information that a participant can use to judge *Emily's* sentence, Li's final prompt goes a step further: she introduces a totally new character, *Kermit*. This character, and not Emily, utters the sentence that the participant has to judge for its truth value. Kermit is «outside of Emily's world» (p. 424): he is part of the linguistic community of the narrator, from whom he learns the vignette

and gains privileged access to the story. Presented below are some excerpts stressing this aspect:

> Probes in the External Condition should be comprehended from the perspective of an omniscient character who is *outside of Emily's world* (e.g. a smart puppet named Kermit)
> (Li, 2021, p. 424, my emphasis)

> […] in this condition [the External Condition], the TVJ questions should be answered from Kermit's perspective, from which the state of affairs is supposed to be *the same as that from participants' viewpoint*.
> (*Ivi*, p. 425, my emphasis)

> Similarly, in the External Condition, since the participants and the puppet Kermit are *in the same epistemic world that is external to Emily*, they all possess privileged knowledge about the events that is not available to Emily and her fellows.
> (*Ivi*, p. 427, my emphasis)

It is important to emphasize an aspect that Li (2011) fails to appropriately address. One may argue that, in light of the above considerations, the descriptivist participant will plausibly judge Kermit's use of "Flint" on the basis of the reference-fixing description that is in force within the narrator's linguistic community, to which Kermit himself belongs. The narrator indeed uses "Flint" to talk about Flint, and according to DT that reference happens through the description "The person whom Newrock's mayor mistakenly took as the hospital builder". If this analysis is correct, also the descriptivist participant, as the causal-historical one, will evaluate Kermit's utterance by taking "Flint" as a name of Flint. Therefore, Li's External Condition seems to prompt the same TVJ from both theories, thereby precluding any comparison between them. The problem at hand, once again, is the possible co-existence of two different reference-fixing descriptions: one in force within Emily's linguistic community and an alternative one in force within the narrator's community, which has Kermit as a member. While discussing the Super Dog Race in Li et al. (2018), in which the speaker is Emily and not Kermit, Li (2023a) explores the possibility that participants approach the TVJ prompts by adopting

the alternative reference-fixing description deriving from their privileged access to the events. She writes:

> It might be argued that the data pattern of the American participants appears to be compatible with a different version of description theory than the one for the Chinese participants mentioned earlier. Such a description theory of reference takes the definite description in the minds of the participants who are the omniscient beings observing Emily's world to be relevant. For these participants, they could associate a different definite description with the dog Pickles, probably to the effect that he was "the dog that was accidentally announced to be the winner of the Super Dog Race" and the description "the dog that won the Super Dog Race" can only be true of the dog Max. If this is the case, contrary to the prediction that a "No" answer to Emily's statement about Pickles is a causal-historical response, people could reject the statement while displaying descriptive intuitions about names.
> (Li, 2023a, p. 61)

Li rejects that, in a vignette like the Super Dog Race, a descriptivist participant may perform the TVJ task by using the alternative, privileged reference-fixing description.

> Yet, notice that such a descriptivist theory of reference posits that reference neither holds in the mind of the speaker who is using the name, nor in the mind of the person who is listening to the speaker or the other people in the linguistic community who might be involved. Instead, it assumes that reference resides in the meta-cognitive awareness of the third party who is observing and judging the use of names by the speaker from a different world that is alien to him/her. Such a conjecture, however, faces […] serious problems. […] it is hard to imagine whether this version of descriptive theory is psychologically real. As reference making is fundamentally about communication and figuring out the referent of names used by speakers, it is unclear how communication could succeed if the speakers and their relevant context were stripped away.
> (Li, 2023a, p. 61)

The speaker who utters the sentence is Emily, who is talking with her father. Both belong to a linguistic community with limited access to the events.[90] Therefore, the only reference-fixing description that can be operative in such a conversation is the one of their linguistic community. According to that description, "Pickles" refers to the real winner of the race, Max.

Li is right that the purported alternative, privileged reference-fixing description cannot be relevant when a participant judges Emily's sentence. However, Li never denies that that alternative reference-fixing description is available and potentially relevant in conversational contexts that do not involve characters like Emily as speakers. Actually, Li seems even obliged to assume that that alternative reference-fixing description is available, because otherwise she could not explain how, from a descriptivist point of view, the narrator can use "Pickles" to talk about Pickles (the dog who crossed the line second) and not Max (the real winner of the race) over the course of the vignette. The External Condition, featuring the introduction of Kermit, precisely provides a context in which the alternative reference-fixing description becomes relevant to the participant's TVJ. Kermit is «outside of Emily's world» and therefore, unlike Emily and her father, Kermit is in the same epistemic position as the participant and is part of the narrator's linguistic community (Li, 2021, p. 424). So probably is also his father, namely the addressee. Within the utterances of the speakers of that community, "Flint" is a name of the individual to whom the mayor mistakenly attributed the construction of the hospital: that is indeed how the narrator uses the name while presenting the vignette. According to DT, that reference relies upon a description like "The individual whom Newrock's mayor mistakenly took as the builder hospital". Therefore, the descriptivist participant will evaluate Kermit's utterance according to that use of the name, and thus judge his sentence as false. Also the causal-historical participant will express an analogous TVJ, and the predictions deriving from the two theories will coincide.

[90] To reiterate: according to DT, Emily and her fellows' perspective on the events is limited not due to their lack of knowledge about Max, the real winner of the race (leaving aside a potential social interpretation of the predicate). They indeed know that Max is the real winner of the race, as the referent of "Pickles" appearing in their belief gets its referent through the description that they associate with the name. However, Emily and her fellows' perspective on the events is limited because they are not privy to some pieces of information to which, conversely, the participant has access thanks to the narrator's account of the events. Specifically, the participant knows that Max felt an overwhelming sense of pride, continued running until he reached the North Pole and subsequently disappeared without anyone seeing him again. Moreover, and crucially, the participant knows that the description that Emily and her fellows associate with "Pickles" does not pick out the original bearer of the name, which entails a divergence between the satisfier of the description itself and the individual at the end of the name's casual-historical chain.

2.3.1.4 Conclusions

In the Original Condition, Li detects a cross-cultural variation in the participants' TVJs. Americans strongly tend to assert that Emily's sentence is false, while Chinese are inclined to assert that it is true. Regardless of the cultural background, those participants who express the TVJ "True" justify it by appealing to Emily's epistemic perspective. Conversely, those participants who provide the TVJ "False" rely on what Li calls the "narrator's epistemic perspective" to substantiate their stance. Concerning the Internal and External Condition, however, the results across the two groups are similar. In the Internal Condition, both Americans and Chinese tend to assert that Emily deserves extra credit. In the External Condition, both Americans and Chinese assert that Kermit's sentence is false.

Figure 2.1 summarizes the results of Li's study:

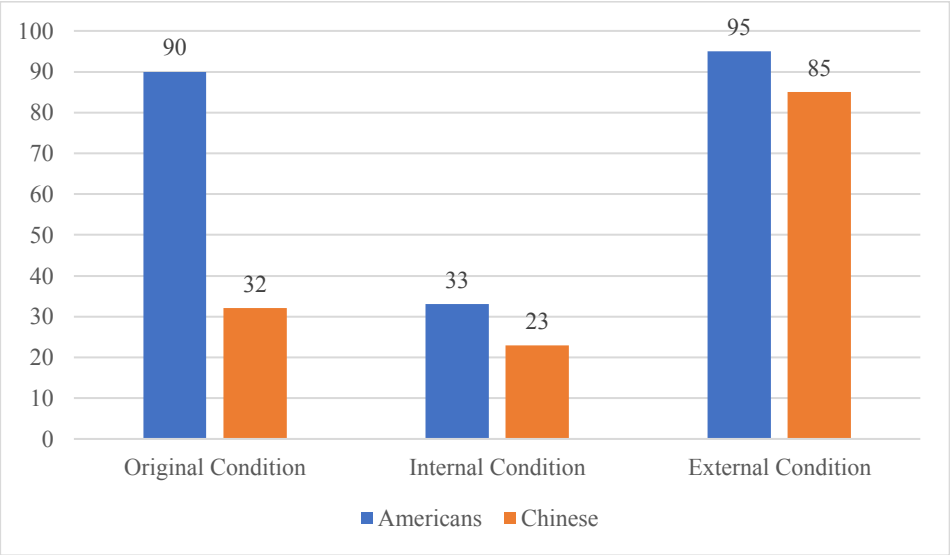

Figure 2.1 Percentages of answers that the speaker is not right or deserves no extra credit in Li's (2021)

Drawing upon these results, Li argues that Chinese participants tend to interpret the TVJ task in the Original Condition as conceptually equivalent to the inquiry of the Internal Condition. Instead, American participants tend to interpret the TVJ task in the Original Condition as conceptually analogous to the query of the External Condition. These different ways of interpreting the TVJ task determine the observed cross-cultural

difference. However, Li argues that neither the Internal nor the External Conditions yield data suitable for testing the theories of reference:

> […] only the original condition in the current study is most likely to elicit judgments that could directly bear on the fundamental issue of reference, whereas the other two conditions may mainly serve the purpose of comparing participants' response patterns under different modes of perspective taking. For in the TVJ tasks in the original condition, judging the truth-value of the critical statements (i.e., Flint was the person who built the hospital) requires participants to figure out the reference of proper names like "Flint" occurring therein. Hence, in theory, the significant differences in American and Chinese speakers' judgments of the statements could be the result of their different views on who "Flint" refers to in the scenarios. But as revealed by the findings […] [concerning the Internal and External condition], these variations are more likely to arise from the differentiated perspective taking tendencies, which thus undermines the conjecture of distinct manners of fixing the reference of names.
>
> (Li, 2021, p. 436)

The Internal Condition is inadequate because it does not address the fundamental point of disagreement between the two theories. DT and CHT hold contrasting views upon whether, in addition to uttering a sentence that reflects what the teacher told her, Emily utters a factually true sentence. As regards the External Condition, Li claims that it fails to provide suitable data to compare the two theories, but she does not explicitly elucidate a precise rationale for this claim. I have argued that the External Condition does not provide adequate data because DT and CHT disagree on the truth value of the sentences that a speaker in semantic contradiction utters. That is not the condition in which Kermit finds himself, though.

In conclusion, the TVJ task in the Original Condition should provide researchers with data suitable to test the two classical theories of reference. However, according to Li, Americans tend to interpret the TVJ prompt as conceptually analogous to the query appearing in the External Condition and Chinese to the query in the Internal Condition.

Problematically, those two models of interpretation are inadequate for the purpose of testing theories of reference.[91]

2.3.2 A critical analysis

2.3.2.1 A comparison with Domaneschi and Vignolo (2020)

As seen, Domaneschi and Vignolo (2020) on the one hand and Li (2021) on the other interpret the so-called "narrator's perspective" differently. As discussed in section 2.2.2.2, Domaneschi and Vignolo use the label "narrator's perspective" to refer to the approach of that participant who judges the *factual truth value* of the sentence *as a speaker in a state of semantic contradiction* utters it. That construal of the narrator's perspective is what Domaneschi and Vignolo try to implement in their Experiment 2. Instead, in Li (2021), the label denotes the evaluation of a sentence as a speaker who is not in semantic contradiction utters it. That construal of the narrator's perspective is what Li reproduces in her External Condition. These differences explain the diverging conclusions that Domaneschi and Vignolo on the one hand and Li on the other draw from their respective data: the former interpret their results as supporting CHT, while the latter as failing to provide any suitable evidence to compare DT and CHT.

As argued, the label "narrator's perspective" is misleading. That said, the correct way of interpreting the so-called "narrator's perspective" is the one that Domaneschi and Vignolo advocate. A classical Gödel-style scenario, such as the one that the Original Condition exemplifies, specifies that the speaker uttering the critical sentence is in semantic contradiction – for instance, Emily. Therefore, even if a participant adopts some non-literal interpretation of the truth predicate, it will likely be one in which the participant is aware that the speaker is Emily. As Domaneschi and Vignolo stress, a participant can potentially engage with the final prompt in two ways: she can evaluate whether what Emily says is factually true or whether reflects what Emily believes. In the former case, the participant understands the truth predicate literally and evaluates the factual truth value of Emily's sentence in light of the privileged information that the narrator provides. In the latter case, the participant departs from the literal interpretation

[91] Li (2021) also conducts another study, Experiment 2, serving as a within-subjects version of the preceding study (Experiment 1), which I have presented in the main text. Experiment 2 involves Chinese participants only. In the Original Condition, participants answer that Emily's claim is true 56% of the time. In the Internal Condition, this figure rises to 88%; in the External Condition, participants assert that Kermit's statement is false in 77% of instances. Therefore, at least among Chinese participants, Li's results are replicable in the context of a within-subjects design.

of the truth predicate and evaluates Emily's utterance in light of the limited information to which Emily has access. Regardless of the approach, though, the participant remains aware that the speaker is Emily; what changes is the (mis)interpretation of the truth predicate.

Instead, Li's External Condition introduces a totally new character, Kermit. In spite of Li's claim to the contrary, the External Condition does not seem to provide a plausible (mis)reading of the Original Condition's final prompt. A participant can take the External Condition's final prompt as a mode of interpretation of the Original Condition's only if the participant, when addressing the Original Condition, *totally* ignores that the speaker is Emily and implicitly replaces her with a new character, like Kermit, who holds an entirely different epistemic position. That replacement is highly radical, given that the Original Condition explicitly indicates that the speaker is Emily and that she belongs to a community with limited access to the events. A theorist needs strong evidence to claim that participants perform such a drastic and abrupt substitution. Li does not offer that evidence, though. The justificatory remarks in the Original Condition after the TVJ "False" that Li provides do not suggest any such replacement, as all the participants' comments mention Emily and the fact that *she* is not right. Moreover, the congruency between the percentages of TVJs "False" from Americans across the Original and the External conditions does not show that Americans understand the Original Condition's final prompt as conceptually analogous to the one of the External Condition. The congruency can simply be due to the fact that the participants regard both sentences (Emily's and Kermit's) as false, as indeed CHT predicts. In no way does that correspondence of TVJs entail that the participants fail to keep track of the fact that the speaker in the Original Condition is Emily. Therefore, absent compelling contrary evidence, the External Condition's final prompt is not a mode in which a participant can plausibly read the Internal Condition's one.

Li (2023b) objects to the above analysis, which I develop in my (2023), in this way:

> According to D'Agruma, if participants in the original condition respond from Emily's perspective as exemplified in the internal condition, then their judgments are "not suitable to test" the two theories of reference, because both theories predict that Emily will receive the extra credit if she utters that name-containing sentence. If they respond from Kermit's perspective as exemplified in the external condition,

> their answers are also useless because both theories "make the same prediction" and hence "an experimental comparison is precluded" [D'Agruma, 2023, p. 1149]. In justifying for this conclusion, D'Agruma appeals to the "New Meaning Objection" originally put forward by Devitt and Porot (2018) which claims "the vignette introduces an additional descriptivist meaning" that is different from the only definite description available to Emily and her cohort [*ibidem*]. But I doubt this meaning will arise, as towards the end of each prompt it is emphasized at least twice that people in the imaginary world have extremely limited information about the historical events and the characters in the stories.
> (Li, 2023b, p. 1161).

It is beyond doubt that, as Li writes, the «people in the imaginary world» have limited access to the events. However, my critical point is precisely that the External Condition, unlike the Original Condition, introduces a speaker, Kermit, *who is not part of that community and who has instead the same epistemic access as the narrator*. Therefore, contingent upon the validity of DT, in Kermit's utterances "Flint" will operate through the same reference-fixing description that the narrator herself uses, namely, a description that picks out Flint – the person whom the mayor mistakenly takes the as the hospital builder – and not Basil – the real hospital builder.

2.3.2.2 The speaker's perspective and the community's perspective

Although Li's External Condition is not suitable to compare DT and CHT, that condition has significant value because it indirectly sheds light on the non-literal interpretation that may arise when participants address the final prompt of the Original Condition. As seen, the Original Condition reproduces a classical Gödel-style scenario, in which the speaker, Emily, is in semantic contradiction. Until now, I have presented the participants who possibly depart from a literal interpretation of the final prompt as adopting the speaker's perspective and judging whether Emily's sentence aligns with the consensus within her community. However, as already mentioned (section 1.4.3.3, n. 44), this formulation is somewhat imprecise, because the perspective of a single individual – as Emily – is different from the perspective or consensus of a group – as Emily's community. What an individual believes does not necessarily align with the broader consensus of the community to which she belongs. The classical-Gödel like scenarios, such as Li's

Original Condition, describe a situation in which Emily and *any other speaker* of her community believes, for example, that "Flint was the person who built the hospital". Therefore, under the supposition that some participants express the TVJ "True" by adopting a non-literal interpretation of the truth predicate, their behavior is consistent with two different explanatory hypotheses. The first is that they express the TVJ "True" because they mean that Emily expresses a sentence that *she herself* takes to be true. The second is that they choose that TVJ because they mean that Emily expresses a sentence that *her community* at large believes, namely a sentence that reflects the community's consensus. However, the two perspectives – the individual's and the community's – constitute different epistemic viewpoints and do not invariably coincide. Consider for example a speaker, John, and his linguistic community, and imagine that John has the belief that "apples are not healthy", while all the other individuals of his community believe that "apples are healthy". In such a circumstance, John's belief diverges from his community's consensus, namely from what his community at large believes.

Although Li does not point out this aspect, her External Condition enables us to ascertain which kind of non-literal interpretation may prompt a participant to choose the TVJ "True" in the Original Condition and thus express an only seemingly descriptivist answer. In the External Condition, Kermit believes that "Flint was the person who built the hospital". Otherwise, he would not utter that sentence, particularly considering the absence of any indications suggesting falsehood or an intent to mislead his addressee – in this case, his father. Therefore, if the participants had any inclination to perform TVJ tasks by adopting the single speaker's perspective, they should answer that Kermit is right (i.e., that his utterance is true), given that Kermit's claim aligns with what he himself believes. However, that is not what Li finds, with both American and Chinese participants. Therefore, if the supposition that some participants adopt a non-literal interpretation of the truth predicate in the Original Condition is correct, then their approach involves the evaluation of whether Emily's statement aligns with the consensus prevailing within her community.

Domaneschi and Vignolo (2020) contrast the «narrator's epistemic perspective» with the «hypothetical character's epistemic perspective», without explicitly delineating the possible distinction between what a single speaker believes and what a community at large does (p. 441). Instead, Li considers both of them. While in (2021) Li's emphasis is

exclusively on the single speaker's perspective, in her (2023a), while discussing the possible origins of the cross-cultural variation in TVJs, Li considers – among others – two hypotheses, which she calls "Initial Cultural Learning" and "Perspective-Taking Strategies". The former is basically a community-consensus hypothesis:

> In the case of reference fixing of proper names in the given scenarios, American[s] […] may focus on what they have learned from the stories. If they understand the stories well and remember (most or all of) the details, they know that the real winner of the Super Dog Race is the dog Max, but not Pickles. Hence, they will insist on their own knowledge and disregard whatever everyone else in the story has been informed of. They thus reject the test statements. On the contrary, the Chinese […] may have been inculcated in the traditional cultural values. Although they have obtained similar knowledge about the stories in the experiment, they nonetheless incline to follow *what is generally and socially accepted* about the characters in the story settings, and hence accept the test statements.
> (Li, 2023a, p. 64, my emphasis)

The "Perspective-Taking Strategies" hypothesis corresponds to the epistemically individualistic explanation that Li (2021) advocates:

> As a result, in fixing the reference of the proper names used by Emily, participants can either choose to model her mental state and take on her [Emily's] epistemic perspective or insist on their own mind. Taking their own perspective, they should judge the test statements to be false. To them, the name "Pickles" picks out the dog Pickles that was originally given the name, irrespective of whatever activities he had participated in or whatever descriptions that were attached to him in relation to those events. Judgments made in this fashion may accord with the causal-historical view of reference. However, if participants could overcome their own perspectives and come to adopt *Emily's viewpoint*, then the statement is right because in Emily's world the dog Pickles has always been associated with the description "the dog that won the Super Dog Race." That is, reference is mediated by the definite description in *Emily's head* which is also shared among the people in her world.
> (Li, 2023a, p. 66, my emphasis)

As mentioned, Li fails to notice that the External Condition can shed light on the choice between the two hypotheses, and she eventually endorses the epistemically individualistic hypothesis rather than the social-consensus one.[92]

Two final considerations are in order. First, one may doubt that talking about the "community's perspective" or "the belief that the community holds" is appropriate, because only single individuals, and not also communities, have a mind and therefore hold things such as beliefs or perspectives about reality. What pertains to a community is only a consensus: the single speakers' beliefs or perspectives on things can possibly shape the community's consensus, which means that most of the people within that community hold the same belief. Therefore, one may object that it is opportune to employ only the label "community's consensus". In this dissertation, I will talk about the "community's perspective" or "consensus" interchangeably, as nothing essential hinges on that distinction. If the reader deems the label "community's perspective" to be less fitting, she may take it as loose talk for "community's consensus".[93]

Second, Devitt (2023) argues that, alongside the community-consensus hypothesis, another explanatory framework can account for the findings across Li's Original and External Condition: the "Blamelessness Hypothesis". Devitt focuses on some explanatory remarks that participants provide in the Original Condition to justify their TVJ "True", such as for example "We cannot blame the innocent people. From Emily's perspective, she is right" (Li, 2021, p. 430). As seen, Emily in the Original Condition and Kermit in the External Condition utter the same sentence. Devitt argues that the Chinese participants in Li (2021) answer that Emily is right while Kermit is not because, despite Emily being factually wrong (as per CHT), she could not have known otherwise as her entire community holds her belief. Therefore, Emily is blameless. Instead, Kermit should have known, since the community to which he belongs has access to the factual truth regarding the events. Therefore, Kermit is blameworthy. As far as I can see, the Blamelessness Hypothesis and a hypothesis based on the community's consensus are difficult to

[92] As the reader may have noticed, Li talks about epistemic perspectives (whether of the single speaker or of the community) as viewpoints that *mediate* reference – that is, depending on the adopted perspective, the name takes on a descriptivist referent or a causal-historical one. As I stress in the upcoming section (2.3.2.3), contrary to her position in the (2021) paper, Li (2023) uses the perspective-based framework to draw conclusions concerning the reference of proper names.

[93] Social epistemology is the branch studying – among other topics – the so-called "group beliefs", namely the purported beliefs that groups may hold in virtue of the single individuals' beliefs. For an overview of the literature, see Goldman and O'Connor (2021).

distinguish because they represent two sides of the same coin. As said, a speaker is blameworthy or blameless depending on whether she could have known otherwise, which in turn is contingent upon the kind of consensus prevailing within the community to which the speaker belongs. Indeed, Devitt himself suggests that, in alternative to his Blameless Hypothesis, «perhaps hypothesis (1) [a social-consensus hypothesis] is right: Chinese think that Emily was right to say that Flint built the hospital, because she was expressing the consensus in her community. But Kermit is not part of that community and so was not right to say this» (p. 9). He then concludes that Li's experiments «support two promising explanations of cultural variation: social conformity and blamelessness» (*ibidem*). In what follows, the focus will be on the social-consensus hypothesis rather than the blamelessness one because the former is conceptually primary over the latter: it is *because* the community shares a certain consensus or perspective about some events *that* a participant can deem a specific speaker as blameworthy or not.[94]

2.3.2.3 Against referential pluralism

As seen, Li (2021) argues that the Americans and Chinese understand the Original Condition's final prompt differently and that both interpretations are irrelevant to testing theories of reference. For that reason, Li refrains from drawing conclusions regarding the correct theory of reference. However, in her subsequent work (2023a), Li presents a different interpretation of the results of her prior (2021) and embraces referential pluralism. This viewpoint posits that both theories – DT and CHT – are right to some extent. Li writes:

> All in all, taking into account the considerable intra-cultural and cross-cultural variation in people's referential intuitions about proper names […], we deem that the monist proposal is not quite appealing. It is rather unlikely that people are universally pure descriptivists or causal theorists, for if this is a truism there should be no sizable variation of any kind. On the contrary, referential pluralism seems to be the more viable option, particularly for those who are determined to develop a substantive

[94] Li (2023b) develops similar considerations: «[c]onvincing as this account [Devitt's Blamelessness Hypothesis] sounds, I find it essentially just about the same as the perspectivalist account I have sketched. If, in determining the truth-value of Emily's and Kermit's name-containing statements, epistemic perspectives are not relevant and the participants do not have to consider the different knowledge status in the first place, how can they possibly judge whether the characters are blameless or not?» (p. 1160). I agree with Li's analysis, albeit with the already specified difference: the target of the perspective-taking strategies is the community's perspective, and not the single speaker's.

> account of reference that could accommodate the variations observed at the individual and group level.
> (Li, 2023a, pp. 146-7)

The two pluralist positions that Li considers are the Ambiguity Hypothesis and the Hybrid Hypothesis. According to the Ambiguity Hypothesis, a proper name has two distinct reference-fixing mechanisms: one causal-historical and the other descriptivist. In situations like the Gödel Case, those mechanisms lead to different referents. Thus, the name "Gödel" can pick out two individuals: one is the person at the end of the causal-historical chain (Gödel), while the other is the satisfier of the associated description (Schmidt). Depending on the conversational context, speakers employ either of those referential mechanisms. Westerners tend to privilege the causal-historical link, while East Asians lean toward the descriptivist one. Proper names are therefore ambiguous. Similar to a term such as "bank", which can denote the edge of the river and the financial institution, "Gödel" has two referential mechanisms.[95] The ambiguity of names becomes manifest only when the descriptivist referential mechanism and the causal-historical one diverge and lead to different individuals. According to the Ambiguity Hypothesis, in the actual world, the name "Gödel" picks out the same individual – Kurt Gödel – through both its causal-historical mechanism and its descriptivist one. Li draws the Ambiguity Hypothesis from Nichols et al. (2016), Tobia et al. (2020) and Haukioja et al. (2021), who apply it to natural-kind terms.[96]

[95] Three considerations are in order as to the example of "bank", when compared to the supposed ambiguity of the name "Gödel". First, by using the example of "bank", which is ambiguous between the edge of a river and the financial institution, I am not suggesting that the purported ambiguity of names is a case of homonymy rather than polysemy (by the way, "bank" is also polysemous, since its economic sense is ambiguous between the financial institution and the building where the financial transactions happen). Li does not take a stance on the matter. More in general, establishing whether the purported ambiguity of names is a case of homonymy or polysemy is not essential for evaluating the correctness of the proposal. Second, "Gödel" would be ambiguous between two referential mechanisms (a causal-historical and a descriptivist one), which is not necessarily the case with "bank": one can accept that both references of "bank" operate through the same mechanism (e.g., a descriptive one). Third, the two referents of "bank" do not coincide: one thing is a financial institution and another thing is the edge of a river. Instead, the two supposed referents of "Gödel" may coincide, when the description associated with the name selects the entity at the end of the name's causal-historical chain (actually, that is even the standard situation, if Gödel-style scenarios are exceptions rather than the norm).

[96] Although Li mentions Tobia et al. (2020) and Haukioja et al. (2021) as works inspiring the Ambiguity Hypothesis that she applies to proper names, some provisos are probably required. Tobia et al. (2020) and Haukioja et al. (2021) provide data suggesting that both the superficial properties and the deep ones that speakers associate with natural-kind terms alternatively determine the extension of those words. That thesis is compatible with, but does not imply, the kind of variation between CHT and DT that Li applies to proper names. The relevance of both superficial properties and deep ones is also compatible with DT only, if the descriptions that a speaker associates with a natural-kind term also include deep properties. That said, it seems fair to say that there is a common element between Tobia et al. and Haukioja et al.'s proposal on the one hand and Li's on the other, and that lies in the supposed ambiguity of the terms under investigation – respectively, natural-kind terms and proper names. Moreover, it is important to specify that Tobia et al.

According to the Hybrid Hypothesis, a proper name has a single referential mechanism relying upon both a causal-historical chain and a description. Most of the time, the two paths align and point toward the same referent. However, there are cases in which they diverge, as John's use of the name "Gödel" illustrates. In such circumstances, a participant can resolve the ambiguity by privileging one referential element over the other. The choice will depend on the conversational context, although in general Westerners privilege the causal-historical resolution of the divergence while East Asian the descriptivist one.[97] Li draws the Hybrid Hypothesis from Genone and Lombrozo (2012) and Devitt and Porter (2021), who apply it to natural-kind terms.[98]

The core of the Hybrid and the Ambiguity Hypothesis is that, in a certain way, both DT and CHT are correct theories of reference for proper names.[99] Differentiating the empirical predictions of these two views and comparing them represents a nontrivial task. However, for the purpose of our discussion, it is not crucial to rigorously distinguish the two hypotheses. Therefore, while my forthcoming analysis mentions only the Ambiguity

---

inscribe the supposed ambiguity of natural-kind terms in the broader framework of the so-called "dual character concepts" (see Knobe, 2022, section I, for a presentation and analysis of those concepts).

[97] Note that the way Li uses the label "referential pluralism" (and also the way I do) differs from the way Mallon et al. (2009) use the label. According to Li, names' referential mechanism would be pluralist in the sense that the very same name incorporates different referential mechanisms (or different referential elements within the same mechanism, in the case of the Hybrid View). Therefore, while the cultural background of the speaker may incline her to prefer one mechanism over the other, both mechanisms remain available to her. Instead, Mallon et al. (2009) use the label "referential pluralism" to refer to the theoretical position that different cultural groups (or different subgroups within the same culture) would have access to only one referential mechanism of the name: roughly speaking, all Westerners would be only Kripkean users of names; all East Asians would be only descriptivist users.

[98] Note that the pluralist position of the Ambiguity or Hybrid Hypothesis differs from the position that Evans (1973) proposes and that some scholars like Reimer (2010) label as "hybrid" – Evans' posthumously published (1982) refines certain aspects of his theory, but in what follows I focus on his (1973). According to Evans, the referent of a name is the dominant causal source of the majority of the information contained in the description that speakers associate with the name. If one accepts that the dominant causal source of the information does not vary contextually, then a name can have just one referent. Evans himself claims that the referent of a name does «not change from occasion to occasion» (p. 202). Therefore, as Genone and Lombrozo (2012) and Nichols et al. (2016) point out, Evans' theory cannot accommodate an intra or cross-cultural variation in the referent that language users assign to a name. Moreover, as Martí (2015) and Devitt and Porter (2021) emphasize, there are valid reasons to doubt that Evans' theory is actually hybrid. What determines the name's referent is the causal connection between a description and an entity X, while the potential satisfactory relation between that description and X is accidental and plays no role in fixing the name's reference. In contrast, the satisfactory relation is the key element of DT. Therefore, there are grounds to deny that Evans' theory is hybrid: Evans' proposal falls within the causal-historical theories of names.

[99] Also the recent contributions by Machery et al. (2023) and Devitt and Porter (2023) challenge a monist position on the reference of natural-kind terms. Instead, Haukioja et al.'s (2023) empirical findings corroborate a monist position, according to which what determines the reference of natural-kind terms are only the deep properties. Although *prima facie* Haukioja et al. (2021) and Haukioja et al. (2023) appear to advocate conflicting positions, pluralism and monism respectively, a closer examination instead suggests a consistency. Haukioja et al. (2021) contend that superficial properties determine the term's extension when it comes to what they call "reverse Twin Earth scenarios", namely cases in which two samples share the same deep structure but display different superficial properties. Instead, as regards the traditional Twin Earth scenarios, Haukioja et al.'s (2021) data corroborate philosophers' traditional intuitions, according to which the kind's deep properties determine the term's extension. Haukioja et al. (2023) collect analogous results with other Twin-Earth cases, without exploring the reverse scenarios. Therefore, as long as the overlapping type of scenario across Haukioja et al. (2021) and Haukioja et al. (2023) is concerned, namely the Twin Earth one, the conclusions of the two papers align.

Hypothesis, as it is the one that Li considers more corroborated (Li, 2023a, pp. 146-7), my criticisms are equally applicable to the Hybrid Hypothesis. Li (2023a) supports the Ambiguity Hypothesis by building upon the results that she collects from the Internal and External conditions. In her view, the different epistemic perspectives *mediate* the two referential mechanisms. If a participant adopts the speaker's perspective, as the Internal Condition instantiates it, the name takes on the descriptivist link; on the other hand, if a participant adopts the narrator's perspective, as the External Condition instantiates it, the name takes on the causal-historical link. Li explains:

> […] we have gained the insight that what straightforwardly determines which kind of referential links should be employed under given circumstances is likely to be the epistemic perspective from which the vignettes and questions are interpreted. […] […] epistemic perspective has been found to be a critical determiner of the linkage between a referring term and its referent in the Gödel case.
> (Li, 2023a, p. 146)

Li (2023b) restates the same point

> What the experimental results reveal is exactly such an alignment of the cross-cultural pattern of referential judgments and perspective-taking strategies. The point of view from which the hypothetical scenarios are to be construed determines to a large extent the way to fix the reference of names in the imaginary speaker's statement.
> (Li, 2023b, pp. 1161-2)

My criticism is straightforward as it relies upon what Li herself argues in her (2021) and on the considerations that I have already developed. DT and CHT do not disagree upon the fact that a speaker in semantic contradiction, like Emily, utters a sentence that she and her community take to be true. That is precisely why the speaker's or community's perspective, which the Internal Condition exemplifies, does not provide adequate data to compare DT and CHT. As regards the so-called "narrator's perspective", Li operationalizes it with the External Condition, where she asks the participant to judge the truth value of a sentence uttered by a speaker, Kermit, who is not in semantic contradiction. As seen, Li (2021) is not explicit on why the External Condition would fail

to provide suitable data to compare DT and CHT. I have argued that the invalidity of the External Condition lies in Kermit's membership in the narrator's linguistic community. Thus, the descriptivist participant will plausibly assume that the pertinent reference-fixing description to judging Kermit's utterance is the one that the narrator employs. Therefore, Li's External Condition too is unsuitable to test theories of reference, as both the descriptivist and the causal-historical interpretation of the name leads to the same referent. In conclusion, if the Internal Condition and the External Condition indeed reflect the participants' actual modes of interpretation of the Original Condition (as Li maintains), any conclusion regarding the referential mechanism of names – whether monist or pluralist – is precluded.

2.3.2.4 The follow-up question is still theoretical

The TVJ "True" in the Original Condition is epistemically ambiguous. By expressing it, a participant can mean that what the speaker says is factually true or that it aligns with the consensus of the community to which the speaker belongs. In Li's Original Condition, both American and Chinese participants who express the TVJ "True" justify this choice by appealing to Emily's epistemic perspective. For easy reference, again reproduced below are the illustrative remarks that Li (2021) provides:

(1) "Emily is right because that is what she learned. The first person in these stories should stop leaving the town forever right after they do something."
(2) "Emily, as a kid, was easily influenced by her surroundings, she said what she was taught."
(3) "We cannot blame the innocent people. From Emily's perspective, she is right."
(Li, 2021, p. 430)

The justificatory comments stress the community's consensus and what Emily believes.[100] These data seem to corroborate that the supposedly descriptivist TVJs derive from participants who adopt the community's perspective, which is not pertinent to testing theories of reference. These data align with the follow-up results in Experiment 1

[100] As discussed (section 2.3.2.2), in the context of Gödel scenarios, there is no gap between the community's consensus and what Emily believes; however, the External Condition shows that what ultimately may guide the participants' departure from the literal interpretation of the truth predicate is the community's consensus.

of Domaneschi and Vignolo (2020). However, in section 2.2.2.5, I have argued that Domaneschi and Vignolo's follow-up question might not be totally reliable in detecting the rationale prompting a participant to express a TVJ. Let us briefly recapitulate the two reasons supporting that claim. First, the follow-up option that is supposed to capture the answer of the genuine descriptivist participant presents a parenthetical remark that is inaccessible to the descriptivist participant, or at least not immediately accessible. Second, asking a participant to answer a follow-up option implies assigning her a theoretical task, thereby reviving Martí's criticism. One may wonder whether and to what extent Li's methodology is susceptible to the same objections.

The main difference between the follow-up question in Domaneschi and Vignolo (2020) and the one in Li (2021) lies in their nature: while the former is a closed question, the latter takes on the format of an open question. Such a methodology provides less easily analyzable data, as coding participants' explanations to discern whether they originate from the adoption of the community's perspective or not may pose challenges. However, Li's open-answer format circumvents the first problem afflicting Domaneschi and Vignolo's follow-up part, namely the potential inaccessibility of the supposedly descriptivist follow-up option. In Li's setup, both the descriptivist and causal-historical participants should find the follow-up question equally accessible: given its open-answer format, the participant can freely formulate a response by using her preferred wording. Hence, the subject will formulate answers consistent with her way of using names. For example, a descriptivist participant can express her genuinely descriptivist justificatory remark with a statement such as "Emily is talking about Basil, the hospital builder".[101]

However, the other criticism that I have raised against Domaneschi and Vignolo (2020), concerning the theoretical nature of their experiment, also extends to Li (2021). Asking participants to justify their TVJ, even in an open-question format, involves soliciting them to theorize on it. As discussed, in MMNS (2004) a participant may choose the descriptivist RI because she finds the descriptivist explanation theoretically intuitive, even if she is a causal-historical user of names. In the context of a TVJ task, a similar and specular phenomenon can take place. A descriptivist participant may express the TVJ

[101] However, since the vignette's wording in Li (2021) essentially follows the one in Li et al. (2018), its formulation inherits the same bias against DT. As discussed, a descriptivist narrator, unless she adopts a *de dicto* mode of expression of Emily's mental representations, could not utter that "Today, people still learn that Flint built the hospital in Newrock, but this is the only thing they have ever heard about Flint. No one has ever heard of Basil".

"True" for genuinely descriptivist reasons, and yet justify her choice with a remark focusing on the community's consensus as she intuitively perceives that explanation to provide the most compelling theoretical framework for the ascriptions of TVJs.

2.3.2.5 How conclusive are the results with American participants?

One can acknowledge that Li's follow-up question in the Original Condition is susceptible to Martí's criticism of over-theorization. However, when focusing on Li's American sample, one can also straightforwardly note that American participants express the TVJ in accordance with CHT 90% of the time. Those TVJs constitute genuine causal-historical evidence because the potential epistemic ambiguity pertains only to the predicate "Truth". The prevalence of the TVJ "False" seems substantial enough to relegate the remaining 10% of TVJs "True" to the category of "experimental noise".[102] The result that 90% of TVJs are causal-historical seems to constitute sufficiently clear support for the conclusion that CHT is the correct theory of reference for Americans. Although Li does not engage in a detailed comparison, there are some plausible reasons that explain why Li (2021) collects a higher percentage of TVJs "False" when compared to other TVJ studies testing Western participants, like MOD (2009), Machery and Stich (2012), Li et al. (2018), Vignolo and Domaneschi (2018), and Domaneschi and Vignolo (2020, 2022).

The first reason is that the predicate of the critical sentence in Li (2021) contains the description that, according to DT, fixes the name's referent – e.g., "Flint was the person who built the hospital". As argued (section 2.2.2.1), that formulation plausibly prevents the causal-historical participant from interpreting the name attributively. Therefore, that feature of the sentence arguably reduces the proportion of TVJs "True" because it inhibits those TVJs that stem from an attributive reading of the name. As seen, the answers based on that understanding do not constitute a threat to CHT, because attributive uses of names are a descriptive phenomenon that the causal-historical theorist already acknowledges. The second reason that explains the higher percentage of TVJs "False" regards the definite descriptions that Li uses, which do not allow for a social interpretation. For example, Li et al. (2018) use the predicates "Being the winner of the Super Dog Race"

[102] «When we look at the effect of our experimental manipulation, it is always against a background of 'noise' caused by random, uncontrollable differences between our conditions», that is, by factors and aspects over which the experimenter lacks full control (Field, 2009, p. 17).

and “Being the discoverer of the Clay Men”, which, as discussed (section 2.2.2.6), admit a social interpretation. If a participant understands the predicate socially, she will provide the TVJ “True” even if she is a causal-historical user of names. The absence of these confounding factors – the attributive understanding of names and the social interpretations of predicates – probably had an impact on the results that Li (2021) collects.

One may think that, as Li (2021) controls for those confounding factors, the clear causal-historical results may constitute the final word on the debate with Western participants, or at least with American ones. However, there are some reasons to be more cautious about that. Li (2023a) conducts an experiment, Study 3, that aims to ascertain the impact of the different epistemic perspectives. The participants – all Americans – are presented with four vignettes, in two blocks: the first includes the Hospital in Newrock and the Little Monk’s Drawing, which are two of the four vignettes that Li (2021) uses and which do not suffer from the social-interpretation problem; the second block includes shortened versions of the Super Dog Race and The Discovery of Claymen, which are the vignettes that Li et al. (2018) use and which suffer from the social-interpretation problem. The final prompt is a classical TVJ question involving a speaker in semantic contradiction, Emily, analogously to what happens in Li et al. (2018) or in Li’s (2021) Original Condition. Li randomly assigns the participants to two groups, differing for a priming task preceding the vignettes. In the *other-condition*, the prime should enhance sensitivity to the other’s perspective and thus increase the inclination to perform the TVJ from Emily’s viewpoint, thereby raising the proportion of TVJs “True”. Instead, the *self-condition* is supposed to enhance the participant’s focus on her own complete and privileged information.[103]

[103] While the details of the primes are not essential for my purposes, I include them for the sake of thoroughness. Participants have to move certain objects on a screen according to the prompts of a fictional instructor. The difference consists in the instructor’s position: while in the *self-condition* the instructor has the same perspective as the participant, to the effect that for instance the instructor’s right side is the participant’s right side, in the *other-condition* the instructor has an opposite perspective compared to participants, to the effect that for instance the instructor’s left side is the participant’s right side. The prompt does not contain only the direction (“left”, “right” and so on), but also the reference to some geometrical shape: for example, “move the item to the right, onto the blue hexagon”. Since this is a priming task, it is important that the participants perform it as expected, ensuring that the prime can have its purported effect upon the participants’ subsequent answers. For that reason, the prime comprises two stages. In the first, the instructions are such that the mention of the shape unambiguously determines the position where the participant has to move the object. Such an unambiguous determination is crucial in the *other-condition*. Suppose for example that the instructor in that condition asks the participant to move the object “to the right”. At that point, the participant has two potential positions towards which she can move the object: either to the instructor’s right or to her own right. However, between those two competitor positions, only one has the figure that the instructor mentions – e.g., the blue hexagon. In that way, the instruction clarifies the perspective that the participant has to adopt. In the second stage, the mention of the

Table 2.1 reports the proportions of causal-historical TVJs.

Table 2.1 Proportion of causal-historical answers in Li's (2023a) Study 3

Condition	Block 1		Block 2	
	The Hospital in Newrock	*The Little Monk's Drawings*	*The Discovery of Claymen*	*The Super Dog Race*
Self-condition	75%	77%	67%	70%
Other-condition	88%	86%	70%	72%

The patterns that the two conditions elicit do not align with Li's expectations. If anything, they show an opposite trend, to the effect that the percentages of TVJs "False" are numerically higher in each story of the *other-condition* than in the corresponding ones of the *self-condition*.[104] Another point is relevant. Let us consider the Hospital in Newrock and the Little Monk's Drawing, namely the scenarios that do not suffer from the potential social interpretation of the predicate. In the *other-condition*, the proportions of causal-historical answers, 88% and 86%, are very close to the 90% that Li (2021) gathers. However, in the *self-condition*, the proportion of TVJs "True", namely the purportedly descriptivist TVJ, reaches as high as 25% among participants in the Hospital in Newrock and 23% in the Little Monk's Drawing. Therefore, Gödel-style vignettes, even when presenting a design that controls for an attributive understanding of the name and a social interpretation of the predicate, happen to elicit a quarter of answers that do not align with the Kripkean prediction. If CHT is the correct theory of reference, the Kripkean theorist needs to advance some explanation for that residual purportedly descriptivist evidence.

Moreover, the thesis that some additional explanation is necessary seems also to emerge from a closer analysis of The Discovery of Claymen and the Super Dog Race. In

shape does not distinguish between the target spot, where the participant has to move the object, and the competitor spot, because both spots have the same figure. Therefore, the participant needs to mandatorily rely upon the spatial instruction ("right", "left", "forward", "backward"), according to the perspective that has been relevant for the previous prompts. If the participant moves the object in the wrong direction, an error message appears and the participant has to move the object again: those corrections are to ensure that the priming task can exert its purported effect.

[104] Investigating the reason for the lack of effect of the priming task is not essential to the arguments within this section. For discussion, see Li (2023a, pp. 109-111).

both conditions (the *self-condition* and *other-condition*), the percentages of participants opting for the TVJ "True" are consistently higher than those of the first block. As argued, the discrepancy can be due to the availability, within the stories of the second block, of a social interpretation of the predicate, which may elicit the TVJ "True" also from causal-historical participants. However, one should be somewhat cautious when evaluating the real impact of the social interpretation of the predicate on the results. At the end of Study 3, participants have to justify their TVJ. Li collects explanations in line with those in her (2021): the participants choosing the TVJ "True" appeal to Emily's epistemic perspective, while the ones opting for the TVJ "False" do not. Interestingly, after the Super Dog Race and The Discovery of Claymen, some participants do explain that they answered "True" in light of a social interpretation of the predicate: for example, they claim that Emily's sentence "Pickles was the winner of the Super Dog Race" is true because the relevant authority, namely the race announcer, declared Pickles the winner. Li reports the number of participants who, in the Super Dog Race and The Discovery of Claymen, justify their TVJ in light of a social interpretation of the predicate (for each specific condition: the *self* and *other-condition*). When transformed into percentages, the proportion of explanations after the TVJ "True" that appeal to a social interpretation of the predicate oscillates between 28% and 43%. While these percentages are surely not negligible, they represent a minority of the total explanations. Naturally, one can question the reliability of the participants' justificatory remarks, like one can doubt the reliability of any explanation that a layperson formulates to justify a TVJ. However, those follow-up answers at least suggest that it may be somewhat rushed to contend that the social interpretation of the predicate is what explains *all* the TVJs "True" that vignettes such as the Super Dog Race or The Discovery of Claymen elicit. Therefore, some other complementary explanation seems necessary. In addition to the data deriving from Li's (2023a) Study 3, the Super Dog Race and The Discovery of Claymen elicit respectively 34% and 36% of TVJs "True" in Li et al. (2018), with American participants, and the Super Dog Race elicits 42% of TVJs "True" with Italian participants in Domaneschi and Vignolo (2020).

In conclusion, the very high percentage of causal-historical TVJs (90%) among Americans in Li (2021) is probably due to the elimination of some confounding factors: first, Li's vignettes refrain from employing predicates susceptible to a social

interpretation; second the sentence that participants have to judge does not make an attributive interpretation of the name available. However, Li's (2021) results are arguably not the last word regarding American or Western participants because two of the four vignettes in Li (2021) are also at the center of Li's (2023a) Study 3, where they proved able to elicit results that are less consistently and overwhelmingly causal-historical. Moreover, plausibly, the availability of a social interpretation cannot account for all the TVJs "True" that vignettes like the Super Dog Race or The Discovery of Claymen proved able to elicit. Therefore, some complementary explanation seems necessary to account for the entirety of the purportedly descriptivist evidence that the above vignettes happen to prompt, which sometimes even reaches 42% (Domaneschi and Vignolo, 2020). The adoption of the community's perspective is a viable candidate; however, one may raise some suspicions about the methodology of asking follow-up questions, due to their theoretical nature.

2.4 Conclusions

It is possible to test linguistic usage in two ways: via linguistic production and via linguistic comprehension. Devitt and Porot test linguistic production by conducting an EP test. In spite of their claims to the contrary, their study is likely subject to the New Meaning Objection, once one frames that criticism as stressing the co-existence of two established reference-fixing descriptions rather than the instantaneous formation of a new reference-fixing description. Devitt and Porot's follow-up experiment effectively addresses the New Meaning Objection and the results that they collect support CHT as the correct theory of reference for proper names.

However, it is important also to explore whether the results concerning the other side of linguistic usage, namely linguistic comprehension, align with CHT as well. Domaneschi and Vignolo (2020) and Li (2021) have the merit of identifying a possible non-literal interpretation of the truth predicate that participants may adopt when performing the TVJ final prompt. However, contrary to what Domaneschi and Vignolo claim, that non-literal interpretation does not rely upon or imply an ambiguity in the truth predicate. Moreover, contrary to what also Li maintains, the adoption of the community's or speaker's perspective contrasts not with the adoption of the so-called "narrator's perspective", but with the evaluation of the sentence's factual truth. The follow-up

methodology that Domaneschi and Vignolo and Li adopt is inherently theoretical, and thus becomes susceptible to an extension of the criticism that Martí (2009) raises against the seminal MMNS (2004). Therefore, it is important to explore other methodologies to investigate whether the adoption of the community's perspective indeed afflicts the TVJs that participants express and so the extent to which the linguistic-comprehension data support CHT as the correct theory of reference. That is what my experimental studies, presented in Chapter 3 and 4, try to do.[105]

[105] Therefore, following Domaneschi and Vignolo (2020) and Li (2021), my experimental studies rely upon the acceptance of both the presented lines of objection against MMNS (2004) (section 1.4 and 1.5). On the one hand, as one strand of criticism asserts, I agree that scholars should test theories of reference against linguistic usage and not referential intuitions. On the other hand, as the other strand of criticism highlights, I also acknowledge the presence of an epistemic ambiguity, a confounding factor that remains even in TVJ final prompts aimed to capture the comprehension side of usage.

3 Experiment 1 and 2: TVJ studies on the Jonah Case

In this section, I present two TVJ studies, Experiment 1 and Experiment 2, which focus on the Jonah Case as an instrument to investigate the degree to which the TVJ "True" in the Gödel Case expresses a real descriptivist answer as opposed to the adoption of the perspective of the speaker's community. Experiment 1, which has an overall explorative nature, involves both Italian and Chinese participants; Experiment 2 develops and improves the design of Experiment 1 and presents the results with regard to an Italian sample. All the materials and results of Experiment 1 and Experiment 2 are accessible on OSF.[106]

3.1 Background

3.1.1 Islam and Baggio (2020)

To ascertain whether the TVJs "True" within the Gödel cases, as the existing literature explores them, express a real descriptivist answer, I draw inspiration from the methodological approach of Islam and Baggio (2020), who combine RIs on the Gödel and Jonah Case to ascertain the impact of the purported speaker/semantic reference ambiguity. While I previously presented their Gödel and Jonah results separately (in section 1.3.1 and 1.5.1, respectively), in this section I describe the rationale behind their idea of combining the two cases. Let us consider Gödel 1, which is MMNS's (2004) vignette involving the name "Gödel", and Jonah 1, which is MMNS's vignette involving the name "Attila". CHT, DT and a speaker-reference interpretation of the cases would predict different response patterns across those Gödel and Jonah cases. According to Islam and Baggio, CHT predicts that "Gödel" refers to Gödel, because the thief is the individual at the end of the communication chain, and that "Attila" has no referent, because no man is the original bearer of that name – the individual at the end of the chain was known as "Raditra".[107] DT would predict that "Gödel" refers to Schmidt, because that man is the real discoverer of the theorem, and "Attila" has no referent, because no individual satisfies the description associated with the name. Finally, a speaker-reference interpretation of names would predict that "Gödel" refers to Schmidt, because John can

[106] At this link: <https://osf.io/fy7hz/?view_only=1f4ee02ee560444ba280d2d3ea9db856>.
[107] As I have already stressed (section 1.5.1) and I am about to emphasize again, I do not share Islam and Baggio's idea that "Atilla" lacks a referent according to CHT.

only intend to talk about the satisfier of the sole description that he associates with the name, and "Attila" refers to Raditra, as the speaker believes that "Attila existed" and therefore she must intend to talk about Raditra, namely a real historical figure.[108] Table 3.1 summarizes Islam and Baggio's predictions:

Table 3.1 Summary of Islam and Baggio's (2020) predictions

Interpretation of the name	**Gödel Case**	**Jonah Case**
CHT	Gödel	No referent
DT	Schmidt	No referent
Speaker reference	Schmidt	Raditra

Interestingly, the combination of the predictions across the two cases enables us to disentangle the three hypotheses. If Islam and Baggio considered only the Gödel or Jonah Case in isolation, a disentanglement would be precluded. For example, the answer that the speaker in the Gödel Case talks about Schmidt is consistent with both the descriptivist hypothesis and the speaker-reference one. As evident from the preceding sections, I disagree with Islam and Baggio's interpretation of the Gödel and Jonah cases. Following Vignolo and Domaneschi (2022), I believe that the notion of speaker reference lacks useful applicability to the Gödel Case; for analogous reasons, I believe that it does not extend to the Jonah Case either. Furthermore, I disagree that CHT, in the Jonah Case, predicts that the name has no reference. From a Kripkean point of view, the essential element to reference is the chain of communication, not the retention of a single orthographic form. While in section 1.5.1 I criticize the use of two orthographic versions of the same name, that criticism is at the experimental level: the use of two orthographic forms arguably introduces an unnecessary level of complexity into the vignette. From a philosophical point of view, though, CHT does imply that "Attila" refers to Raditra, that is, that "Attila" and "Raditra" are two versions of the same name. While I could raise other reservations regarding the specifics of their interpretation of the cases, in this context the salient factor is the methodology that Islam and Baggio develop, which I consider valuable and thus adopt for my TVJ study.

[108] Again, as I am about to stress, I do not share Islam and Baggio's application of the notion of speaker's reference to the Gödel and Jonah cases.

3.1.2 The epistemic ambiguity in the Jonah Case

Let us redirect our attention to the primary objective, which is to explore the extent to which the TVJ results in the literature on the Gödel Case provide genuinely descriptivist evidence. To this end, the Jonah Case can become paramount, as DT and CHT predict inverted patterns of TVJs across the two scenarios. Given a Gödel scenario in which John utters a sentence like "Gödel is the discoverer of the incompleteness theorem", the sentence is true according to DT and false according to CHT. Let us now consider a Jonah scenario in which John utters the sentence "N exists/existed" – where N is the proper name around which the case revolves. That sentence is true according to CHT because there exists/existed an individual at the far end of N's causal-historical chain. Conversely, the sentence is false according to DT because there is/was no entity satisfying the description that the speaker associates with N. Therefore, the predictions of DT and CHT in the Jonah Case exhibit an inverse pattern when compared to those in the Gödel Case.[109] Also in the context of the Jonah Case it is possible to interpret the predicate "True" from the community's epistemic perspective. A participant can answer that John's utterance is true, meaning that it reflects the shared consensus within the speaker's linguistic community. In this sense, two points are worth emphasizing.

First, in section 2.2.2.3 I have argued that, in the context of the Gödel Case, the adoption of the community's perspective makes sense only for a causal-historical participant. Symmetrically, in the context of the Jonah Case, the adoption of the community's perspective makes sense only for a descriptivist participant: according to DT, the sentence "N exists/existed" is factually false, and thus the descriptivist participant may be inclined to adopt the community's perspective to charitably judge the sentence as "true". Contrarily, according to CHT, the sentence is not only one that the speaker's community takes to be true, but is also factually true. Therefore, the causal-historical participant has no reason to diverge from the standard interpretation of the truth predicate, which centers upon the sentence's factual truth value.

[109] MMNS (2004) themselves highlight the different predictions deriving from DT and CHT as regards a sentence such as "N exists/existed", when they write that, in the context of the Jonah Case, the descriptivist prediction is that «the statement "Jonah exists" is false (given that the name has no referent)» (p. B4, n. 5). On the contrary, the causal-historical intuition is that «Jonah might have existed, whether or not the description is satisfied» (p. B4, n. 6).

Second, the adoption of the community's perspective in the Jonah Case is somewhat more convoluted than in the Gödel one. While it is beyond doubt that John's linguistic community believes that "Gödel is the discoverer of the incompleteness theorem", as that discovery is the only piece of information that speakers associate with the name, it may be slightly more controversial that, in the context of a Jonah Case, John's linguistic community believes that "N existed". For example, let us suppose that John's linguistic community associates the name with a description outlining a very unrealistic entity, such as a goblin. If that is the case and DT is the correct theory of names, John will implausibly believe and thus utter that "N existed", since he likely believes that things such as goblins do not inhabit our world. This example shows that, in the context of a Jonah scenario in which John associates N with the description D, the belief "N exists/existed" is a potential *consequence* of John's association of D with N. The inference operates under the prerequisite that D outlines a realistic entity, whose existence does not appear as an object of dispute to John and his community.

One may wonder why not construct a Jonah Case in which, given the name N and the empty reference-fixing description D, the speaker utters "N is D", along the lines of those Gödel cases in which the speaker utters "Gödel is the discoverer of the incompleteness theorem" or "Pickles is the winner of the Super Dog Race". Such a formulation raises an intricate issue, though: whether a descriptivist participant should judge the sentence as true or not. Literally speaking, according to DT, the sentence "N is D" is not true, because – given N's emptiness – there is no entity to which the sentence can be truthfully attributing the property of being D. However, since D is N's reference-fixing description, a descriptivist participant may charitably take John – i.e., the speaker – as engaging with the enterprise of *explaining* N's reference-fixing conditions, without any ontological commitment. In other words, the participant may assume that an operator like "If he/she/it exists/existed" implicitly precedes the sentence "N is D", in which case John's utterance turns out to be true. An additional conundrum lies in the fact that, if the participant engages in an evaluation of the literal truth value of John's sentence, the predictions deriving from DT and CHT coincide. Both theories agree that "N is D" is not true, albeit for different reasons: according to CHT, the sentence is false because the real entity that

N denotes does not fulfill D; according to DT, N fails to denote any entity and thus the sentence is not true.[110]

If the above points are right, a properly designed Jonah Case requires a context in which the fictional speaker (e.g., John) is disposed to truly believe and utter a sentence of the form "N exists/existed". As seen, in such a circumstance, the TVJ "False" can express only a descriptivist answer. Indeed, if a participant provides an answer that is causal-historical or centered upon the community's epistemic perspective, she will provide the TVJ "True". In other words, in the Jonah Case the TVJ "False" isolates answers that can only express a descriptivist commitment, similarly to how the TVJ "False" in the Gödel Case delimits answers that can only voice a causal-historical stance. If the TVJs "True" that the Gödel Case elicits voice a genuine descriptivist commitment, the Jonah Case should prompt participants to choose the TVJ "False", as that answer, in the context of the Jonah Case, can only voice a descriptivist commitment.

By relying upon the Jonah Case, my studies also serve to fill a significant gap in the existing literature. While scholars have focused on the Gödel Case, Machery (2021) explains that helpful insights can derive from a further investigation of the Jonah Case: «further work is needed to determine whether judgments about reference vary cross-culturally when they are elicited by a Jonah case, and how answers to the Jonah and Gödel cases are related» (p. 540). Li (2023a) expresses a similar remark: «researchers need to expand the scope of the investigation by looking into the Jonah case and other philosophically interesting scenarios that involve proper names» (p. 157).

3.2 Experiment 1

3.2.1 Hypotheses

Experiment 1, which I conducted in collaboration with Litman Huang, tests some TVJ Jonah scenarios across Italian and Chinese participants. The body of literature on the Gödel Case shows that Italians and more in general Westerners tend to endorse the TVJ "False". However, a non-negligible minority of participants opt for the TVJ "True", namely the purportedly descriptivist answer. Even vignettes like Li's (2021) The Hospital in Newrock and The Little Monk's Drawings, which control for the possible attributive

[110] I use the neutral expression "not true" because a descriptivist participant may also regard "N is D" as neither true nor false, given N's emptiness. This problem introduces an extra layer of complexity with the sentence "N is D", which does not emerge with "N existed": according to DT, this sentence is simply false.

construal of the name and the social interpretation of the predicate, sometimes happen to elicit a proportion of TVJs in line with DT as high as 25%. Moreover, the scenarios that Li et al. (2018) introduce, namely the Super Dog Race or The Discovery of Claymen (the former of which is also at the center of Domaneschi and Vignolo, 2020), prompt a proportion of purportedly descriptivist TVJs that a social interpretation of the predicate arguably cannot wholly address. As regards Chinese participants, the literature shows that they are consistently inclined to provide the TVJ "True" in the Gödel cases, even when one controls for the above confounding factors: in Li (2021) the Chinese sample opts for the TVJ "True" 68% of the time. As argued, two explanations arise to account for the proportions of TVJs "True" that the Gödel-style scenarios elicit. The first is that the participants choosing the TVJ "True" express a genuinely descriptivist answer. The second possible explanation is that they provide their answer by adopting the community's perspective, which is unsuitable to compare DT and CHT.

The rationale of Experiment 1 is this: the Jonah Case can isolate those TVJs (if any) that have a genuine descriptivist nature. If a participant expresses a causal-historical TVJ, she will provide the TVJ "True" and likewise if she adopts the community's perspective. The TVJ "False" can stem only from a participant expressing a genuine descriptivist stance. Therefore, if the cross-cultural variation in the Gödel Case has a semantic nature, then the Jonah Case should elicit a higher proportion of TVJs "False" within an East Asian sample compared to a Western one. More explicitly:

> **Hypothesis 1**
>
> Let us suppose that the participants who in the literature express the TVJ "True" in the Gödel Case do so out of a genuinely descriptivist understanding of the name. In that case, under the assumption that they similarly express such a descriptivist commitment in the Jonah Case, one should expect that the Jonah Case elicits a non-negligible minority of TVJs "False" within the Italian sample and a majority of them within the Chinese one, because in the Jonah Case the TVJ "False" can only express a genuine descriptivist commitment.

Hypothesis 2

Let us suppose that the participants who express the TVJ "True" in the Gödel Case do so because they adopt the community's perspective. In that case, one should expect that the Jonah Case elicits minimal or no instances of the TVJ "False" in both the Italian and Chinese sample, because in the Jonah Case the TVJ "False" can only express a genuine descriptivist commitment.

Similarly to some previous works (Domaneschi and Vignolo 2020; Li, 2021; Li, 2023a), this study unfolds in two distinct phases: an initial TVJ one and a subsequent phase wherein participants justify their previous TVJ "True" or "False". While Hypothesis 1 and 2 are confirmed or disconfirmed by the TVJs, which thus represent the core of the experiment, we initially thought that collecting supplementary follow-up data could be useful to investigate whether those align with the TVJ indications. However, as discussed in the previous chapters, follow-up data have a theoretical nature; moreover, the formulation of those follow-up options proved to have some potential flaws that complicate the interpretation of them. Therefore, in this section, I cover the TVJ part of the experiment and the related results, while, in the Appendix, I present the second phase of the experiment.

3.2.2 Material and procedures

We presented six Jonah cases to participants from Italy and Hong Kong. Three of them employ Italian or Western names (e.g., "Ambiorix") and 3 of them Chinese names (e.g., "Qian Yingcai"). Therefore, each Italian/Chinese participant receives 3 targets with Western names and 3 targets with Chinese names. Italian participants receive the vignettes in Italian, while Chinese ones in traditional Chinese. Reproduced below is one of the vignettes, translated into English (the vignette follows Devitt and Port's (2018) Jonah scenario):

Ambiorix Case

Ambiorix was a much-appreciated peacetime administrator of a small area of Gaul. Ambiorix's friends admired him a lot. So, they came up with a tale according to which Ambiorix was a glorious warrior who conducted the

Gauls against the Romans when the Romans attacked Gaul. In fact, no one ever conducted the Gauls against the Romans. The Gauls were taken by surprise and no one conducted their military response. However, the idea that Ambiorix conducted the Gaul army against the Romans is everything that was handed down about Ambiorix through the centuries. Therefore, that idea is everything that today is associated with the name "Ambiorix".

Leonardo is a history student who carefully attended the lessons on Roman invasions. Like any other person, the only description that Leonardo associates with the name "Ambiorix" is "The warrior who conducted the Gaul army when the Romans attacked Gaul".

When Leonardo says «Ambiorix existed», what Leonardo says is
- True
- False
- I don't know

As discussed, the answer "False" can only express a genuinely causal-historical stance. The prompt also presents the option "I don't know" to minimize the number of participants trying to guess. Regrettably, the vignettes incorporate a bias against DT, because – as discussed – a descriptivist narrator could not utter that certain tales are "everything that was handed down *about Ambiorix*". This flaw is corrected in Experiment 2, as I will show.

Throughout the questionnaire, participants receive also 6 fillers. At the end of the filler, the participant has to formulate two TVJs. Since the filler TVJ tasks have to admit only one answer as unambiguously correct, the filler does not introduce any speaker in semantic contradiction and simply asks the participant to judge the sentences "based on what you have just read". The fillers also function as comprehension and attention checks, since they control that the subject pays attention to the sequence of vignettes over the course of the experiment. Here is an example of one of the fillers:

Filler 1

For several years, Xiao Huang was the pet dog of a famous Chinese painter from the 20th century named “Ding Yuanyuan”. Some of Ding Yuanyuan’s paintings depict Xiao Huang. Ding Yuanyuan was very fond of Xiao Huang to the extent that, when the animal passed away, Ding Yuanyuan decided not to have any other pets. In fact, Ding Yuanyuan believed that no one could replace Xiao Huang.

Consider the phrase “Xiao Huang was Ding Yuanyuan’s companion animal for several years”. Based on what you have read, this phrase is:

- True
- False
- I don’t know

Consider the phrase “Xiao Huang was a cat”. Based on what you have read, this phrase is:

- True
- False
- I don’t know

While the final prompt provides the “I don’t know” option, in line with the available response choices for the target questions, the correct answer to one question of each filler is “True” and to the other is “False”. That is, the vignette is always detailed enough to allow participants to avoid the “I don’t know” answer. For example, in this case, the TVJ for the first sentence is “True”, while the TVJ for the second sentence is “False”.

We administered the experiment online, using PsyToolKit (Stoet, 2010, 2017). Following an introductory screen, participants are asked to provide their consent to participate in the experiment. They are presented with the twelve items (the six targets and the six fillers) in random order. The TVJ final prompt at the end of the target or fillers is on the same page as the item itself. The follow-up questions (which are asked only for the targets) are on a separate page, after the participant chooses the TVJ “True” or “False”, while no follow-up question appears for the answer “I don’t know”. After all the items, participants answer a demographic questionnaire. Finally, the participant needs to

further confirm that she wants to send her answer. Absent that additional consent, the system automatically excludes the answer.

3.2.3 Participants

The recruitment of participants for this study was conducted online, using various social network including Facebook, Twitter[111], and other social media platforms. 91 participants completed the Italian version of the questionnaire and 84 participants completed the Chinese version. The participants included in the final analysis meet the following criteria.

(a) They are native Italian/Cantonese speakers.
(b) They live in Italy/Hong Kong.
(c) They spent most of their life in Italy/Hong Kong.
(d) They answered at least 66% of fillers' questions correctly.
(e) They are not philosophers. We count a participant as a philosopher if she has a degree (Bachelor or Master) in Philosophy, she is a University student majoring in Philosophy, she attended three or more university courses in Philosophy or a course in Philosophy of Language.
(f) They did not complete surveys similar to this recently.

23 (25.27%) of the participants taking the Italian version of the survey do not meet one or more inclusion criteria. 1 participant (1.10%) is not a native speaker of Italian. 4 participants (4.40%) do not live in Italy. 4 participants (4.40%) answered less than 66% of the fillers' questions correctly. 9 participants (9.89%) are classified as philosophers. 6 participants (6.59%) took surveys similar to this recently. Therefore, 68 Italian participants are included in the final analysis (M = 31.03; SD = 14.19; 40 females, 28 males) and took an average of 20.72 minutes to complete the test, with a median time of 16.50 minutes.

39 (46.43%) of the participants taking the Chinese version of the survey do not meet one or more of the inclusion criteria. 1 (1.19%) participant is not a native speaker of Cantonese. 11 participants (13.10%) do not live in Hong Kong. 3 participants (3.57%)

[111] Currently rebranded as "X".

did not spend most of their life in Hong Kong. 15 participants (17.86%) answered less than 66% of the filler questions correctly. 14 participants (16.67%) are classified as philosophers. Additionally, from the participants who met the previously mentioned criteria, we excluded three participants due to their excessively long duration taken to complete the questionnaire reported by the survey platform (248, 671, and 1257 minutes). Those prolonged times indicate that the questionnaire was not completed in a single session, but rather across widely separated intervals. The survey platform's duration count captured the total timespan, including these intervals. Therefore, 45 Chinese participants were included in the final analysis (M = 32.38; SD = 8.65; 21 females, 22 males, 2 prefer not to specify) and the took an average of 17.64 minutes to complete the test, with a median time of 13 minutes.

3.2.4 Results

Figure 3.1 reports the TVJ results aggregating the six targets, across the two samples.

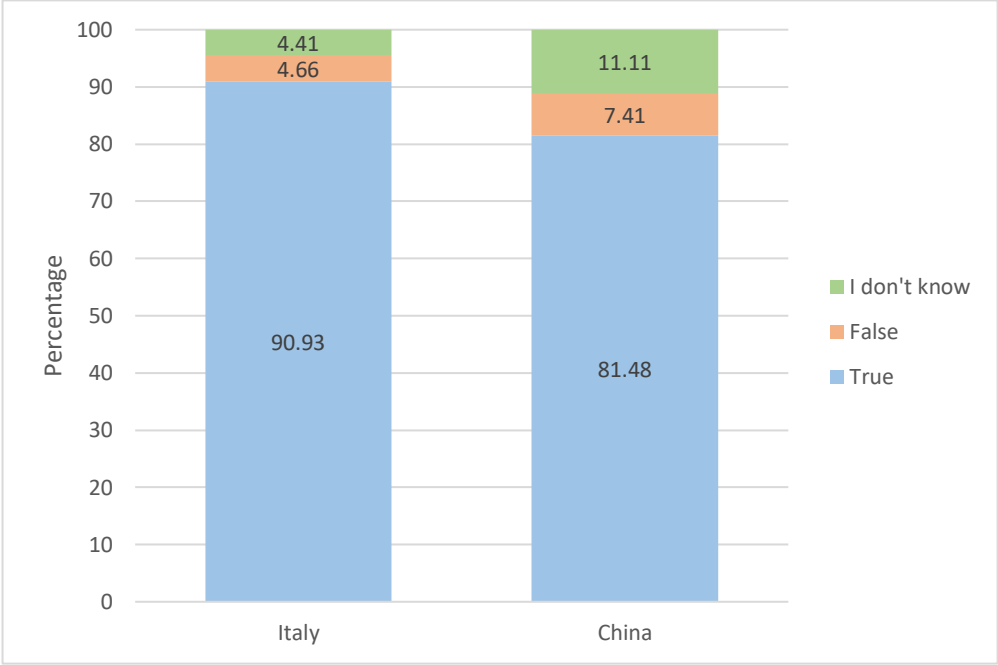

Figure 3.1 Percentages TVJs of Experiment 1

Within the Italian sample, participants choose the TVJ "True" 90.93% of the time, while the TVJ "False" and "I don't know" respectively 4.66% and 4.41% of the time. The Chi-square statistics confirm that the difference in frequency between "True" and "False"

is significant (χ^2 (1, N = 390) = 317.70, p < .001). So is the difference between "True" and "I don't know" (χ^2 (1, N = 389) = 320.33, p < .001). The difference between "False" and "I don't know" is not significant (p > .05).[112]

Within the Chinese sample, participants choose the TVJ "True" 81.48% of the time, while the TVJ "False" and "I don't know" respectively 7.41% and 11.11% of the time. The Chi-square statistics confirm that the difference in frequency between "True" and "False" is significant (χ^2 (1, N = 240) = 166.67, p < .001). So is the difference between "True" and "I don't know" (χ^2 (1, N = 250) = 144.40, p < .001). The difference between "False" and "I don't know" is not statistically significant (p > .05).

To determine whether the frequencies differ across the two samples, a multinomial logistic mixed effects model with culture as fixed effect and participant and vignette as random effects was constructed. The effect of culture is not statistically significant, although it approximates significance (F (2, 674) = 2.54, p = .079). The fixed coefficients of the pairwise comparisons of the multinomial model have "True" as the baseline category. The comparisons reveal that, while Italians are significantly less inclined to choose the option "I don't know" than "True" when compared to Chinese (ß = -1.09, SE = .52, t = -2.09, p < .05, Exp(ß) = .34), no statistically significant difference emerges when the option "True" is compared with "False" (ß = -0.44, SE = .47, t = -.92, p > .05, Exp(ß) = .65).[113]

Table 3.2 and 3.3 report the percentages across the six targets for the Italian and Chinese sample.

Table 3.2 Percentages of TVJs for the six items in Experiment 1 – Italian sample

	Ambiorix Case	Anaximenes Case	Collaborator Case	Fabrication Case	Murder Case	Sister Case
True	86.76	86.76	92.65	95.59	92.65	91.18
False	7.35	2.94	4.41	1.47	2.94	8.82
I don't know	5.88	10.29	2.94	2.94	4.41	0

[112] All the statistical analyses for Experiment 1 and 2 were conducted using SPSS 29.

[113] In the initial phase of our analysis, the model also included a random effect for the slope of culture across vignettes, but we later removed this effect. Its variance in the pairwise comparison between "True" and "False" was 0, leading to a non-positive-definite final Hessian matrix of the multinomial model, which may make the validity of the model fit uncertain. That said, the statistics of the model including the slope are virtually the same as the one not including it. Omnibus test: F (2, 674) = 2.38, p = .093; "True" vs. "I don't know": ß = -1.10, SE = .55, t = -2.00, p < .05, Exp(ß) = .33; "True" vs. "False": ß = -.43, SE = .47, t = -.92, p > .05, Exp(ß) = .65).

Table 3.3 Percentages of TVJs for the six items in Experiment 1 – Chinese sample

	Ambiorix Case	Anaximenes Case	Collaborator Case	Fabrication Case	Murder Case	Sister Case
True	84.44	77.78	84.44	80.00	82.22	80.00
False	6.67	11.11	2.22	6.67	2.22	15.56
I don't know	8.89	11.11	13.33	13.33	15.56	4.44

The percentages across the six vignettes for the Italian and Chinese samples overall align with the descriptive statistics compiled from the six items – no single item's results markedly deviate from the aggregated statistics. Fisher-Freeman-Halton's exact tests[114] confirm that the differences in frequency between the two groups for each single item are not significant, with the exception of the Fabrication Case ($p < .05$). A contributing factor to the statistical significance of the Fabrication Case is also the frequency of responses indicating "I don't know". More in general, the Chinese participants are more inclined to choose that option than Italians. However, the option "I don't know" serves the purpose of reducing the tendency of participants to make hasty or random guesses, and therefore is philosophically inconsequential: what matters is the comparison between "True" and "False". As seen, while the pairwise comparison of the multinomial model between "True" and "I don't know" across the two cultural groups is statistically significant, the comparison between "True" and "False" does not reveal any statistically significant

[114] When the Chi-square statistics assumption that all the cells' expected values are greater than 5 is violated, this dissertation reports the Fisher's statistics (Field, 2009, p. 692; Howitt and Cramer, 2014, p. 205). The Fisher-Freeman-Halton's exact test is an extension of the Fisher's test for contingency tables that exceed a 2×2 format.

difference. Thus, Figure 3.2 presents a clearer depiction of the results by excluding the "I don't know" responses.

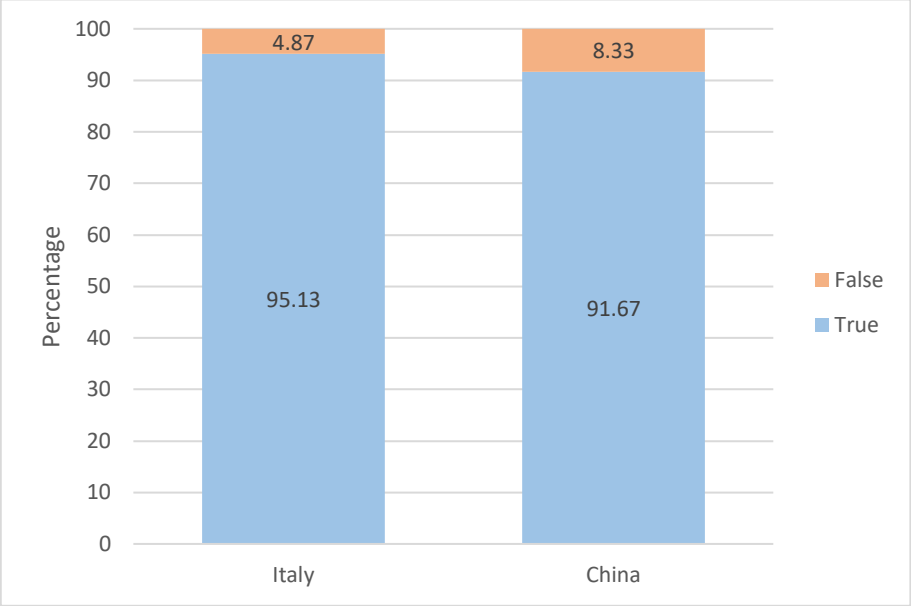

Figure 3.2 Percentages of TVJs with only "True" and "False" in Experiment 1

Table 3.4 and 3.5 report the percentages across the six targets for the Italian and Chinese samples by excluding the "I don't know responses".

Table 3.4 Percentages of TVJs "True" / "False" for the six items in Experiment 1 – Italian sample

	Ambiorix Case	Anaximenes Case	Collaborator Case	Fabrication Case	Murder Case	Sister Case
True	92.19	96.72	95.45	98.48	96.92	91.18
False	7.81	3.28	4.55	1.52	3.08	8.82

Table 3.5 Percentages of TVJs "True" and "False" for the six items in Experiment 1 – Chinese sample

	Ambiorix Case	Anaximenes Case	Collaborator Case	Fabrication Case	Murder Case	Sister Case
True	92.68	87.50	97.44	92.31	97.37	83.72
False	7.32	12.50	2.56	7.69	2.63	16.28

Fisher's tests indicate that cultural background has no significant effect on any of the individual items (*p*s > 0.05), including the Fabrication Case.[115]

3.2.5 Discussion

Participants almost uniformly select "True". If we exclude the "I don't know" option, Italian participants opt for the TVJ "True" 95% of the time, while Chinese participants choose it 92% of the time. These TVJ results corroborate the thesis that there is no genuine descriptivist evidence for the reference of proper names. Rather, the cross-cultural variation that the Gödel scenarios unveil does not have a semantic root but rather stems from a higher tendency among Chinese to adopt the community's perspective. The Jonah Case, where the TVJ that DT predicts deviates from the one corresponding to the adoption of the community's perspective, elicits limited occurrences of the descriptivist TVJ ("False") within both cultural groups.

[115] For the sake of completeness, this note details the results from all the participants who completed the questionnaire, including those who do not meet the inclusion criteria. 91 participants took the Italian version of the questionnaire (M = 30.90, SD = 13.45, 53 females, 38 males); 84 participants took the Chinese version of the questionnaire (M = 33.17, SD = 9.15, 41 females, 40 males, 3 prefer not to specify). Within the Italian sample, the percentages of TVJs are these: "True" 87.00% of the time, "False" 4.95%, "I don't know" 8.06%. The frequency of "True" differs statistically significantly from "False" (χ^2 (1, N = 502) = 399.81, $p < .001$) and "I don't know" (χ^2 (1, N = 519) = 357.92, $p < .001$), and the frequency of "False" from "I don't know" (χ^2 (1, N = 71) = 4.07, $p < .05$). Within the Chinese sample, the percentages of TVJs are these: "True" 79.56% of the time, "False" 7.94%, "I don't know" 12.50%. The frequency of "True" differs statistically significantly from "False" (χ^2 (1, N = 441) = 295.51, $p < .001$) and "I don't know" (χ^2 (1, N = 464) = 246.22, $p < .001$), and the frequency of "False" from "I don't know" (χ^2 (1, N = 103) = 5.14, $p < .05$). A multinomial logistic mixed effects model, with culture as fixed effect and participant and vignette as random effects, reveals that the cultural background does not influence the TVJ choice (F (2, 1046) = 2.44, $p = .088$). The fixed coefficients of the pairwise comparisons of the multinomial model have "True" as the baseline category. The comparisons reveal that no statistically significant difference emerges when the option "True" is compared with "I don't know" (ß = -.72, SE = .39, t = -1.84, $p = .066$, Exp(ß) = .49) or "False" (ß = -0.47, SE = .36, t = -1.29, $p > .05$, Exp(ß) = .63). (The model omits the random slope for culture across vignettes because otherwise the model would have had a non-positive-definite Hessian matrix; anyhow, a model including that random effect leads to the same non-significant results.) Table 3.6 reports the percentages across the six targets for the Italian and Chinese sample.

Table 3.6 Percentages of TVJs for the six items in Experiment 1. ■ Italian sample ■ Chinese sample

	Ambiorix Case		Anaximenes Case		Collaborator Case		Fabrication Case		Murder Case		Sister Case	
True	84.62	80.95	81.32	75.00	87.91	84.52	91.21	80.95	90.11	79.76	86.81	76.19
False	6.59	5.95	4.40	11.90	3.30	4.76	4.40	7.14	2.20	4.76	8.79	13.10
I don't know	8.79	13.10	14.29	13.10	8.79	10.71	4.40	11.90	7.69	15.48	4.40	10.71

Fisher-Freeman-Halton's exact tests and Chi-square tests, selected based on their appropriateness for the data, consistently show no significant differences in frequency between the two groups for any item (all *p*s > 0.05). If one rules out the option "I don't know", the percentages for the TVJ "True" and "False" in the Italian sample are 94.62% and 5.38%, respectively; for the Chinese sample, these percentages are 90.93% and 9.07%. Also in this case, Fisher-Freeman-Halton's exact tests and Chi-square tests, selected based on their appropriateness for the data, consistently show no significant differences in frequency between the two groups for any item (all *p*s > 0.05). Thus, the results from the full group of survey respondents overall align with those from the sample of the final analysis.

3.3 Experiment 2

3.3.1 Hypothesis

Experiment 2 has a design similar to Experiment 1, but constitutes a methodological improvement under three aspects. First, each participant receives both some Gödel scenarios and some Jonah ones. This design enables us to examine the linguistic behavior of the very same participant across the two kinds of vignettes. Testing only the Jonah scenarios while relying upon the Gödel results of the previous literature yields less informative and more speculative insights about the single participants' patterns of answers. Second, I eliminated the follow-up questions: the primary advantage of combining the Jonah TVJ results with the Gödel ones (either from prior literature or directly tested) is precisely the possibility to control for the epistemic ambiguity without depending on theoretical follow-up answers. Moreover, this approach reduces the cognitive effort demanded of the participant and keeps TVJs close to linguistic usage. Indeed, if the participant is asked to answer a follow-up question after a TVJ, that design may lead her to engage in meta-theoretical reasoning not only when she provides the follow-up answer itself, but also when she formulates the TVJ responses to the subsequent vignettes. In other words, the follow-up answers may incline participants to express their subsequent TVJs not out of the propositions that they understand as expressed by the sentences, but out of theoretical reflections. Third, I have improved the wording of the Jonah scenarios, by avoiding any bias against DT. While Experiment 2 involves only Italian participants, the hypotheses are structurally analogous to those of Experiment 1:

> **Hypothesis 1**
>
> Let us suppose that the participants who express the TVJ "True" in the Gödel Case do so out of a genuinely descriptivist understanding of the name. In that case, under the assumption that they similarly express such a descriptivist commitment in the Jonah Case, one should expect that the Jonah Case elicits an analogous amount of TVJs "False" within the Italian sample, because in the Jonah Case the TVJ "False" can only express a genuine descriptivist commitment.

Hypothesis 2

Let us suppose that the participants who express the TVJ "True" in the Gödel Case do so because they adopt the community's perspective. In that case, one should expect that the Jonah Case elicits minimal or no instances of the TVJ "False", because in the Jonah Case the TVJ "False" can only express a genuine descriptivist commitment.

3.3.2 Material

3.3.2.1 Gödel-style scenarios

I presented 2 Gödel-style cases and 2 Jonah-style cases in Italian to Italian participants. These are the two Gödel-style vignettes:

Mathematician Case

Heumerus was an ancient Greek mathematician who discovered an important theorem, known as the "prime number theorem". However, Heumerus died shortly after making that discovery, without being able to make it known. Another mathematician, Theodorus, stole the manuscript containing the proof. Theodorus then presented himself as the discoverer of the prime number theorem. No one ever found out about the theft. As a result, the statement "Theodorus was the discoverer of the prime number theorem" is what everyone believes today; furthermore, this statement is all that has been handed down over time in relation to the name "Theodorus".

Marta is a student who took a history of mathematics course, during which she also studied the prime number theorem. Like everyone else, the only description that Marta associates with the name "Theodorus" is "The discoverer of the prime number theorem".

When Marta says «Theodorus was the discoverer of the prime number theorem», what she says is

- True
- False

Painter Case

Sara Verdi was a painter from the early 18th century and was the author of a painting that is now exhibited in Milan, titled *Happiness*. Sara Verdi gave her painting to an exhibition organizer so that he could display it during various exhibitions. However, the organizer mixed up the authors of some paintings and, specifically, thought that the author of *Happiness* was Elisabetta Rossi, another painter active at the time. As a result, every time the organizer exhibited *Happiness*, he presented it as a painting by Elisabetta Rossi. Since neither Sara Verdi nor Elisabetta Rossi were present on those occasions, they could never discover the mistake and correct it. As a result, the statement "Elisabetta Rossi was the author of *Happiness*" is what everyone believes today; furthermore, this statement is all that has been passed down over time in relation to the name "Elisabetta Rossi".

Gianni is a student who took an art history course, during which he also studied the painting *Happiness*. Like everyone else, the only description that Gianni associates with the name "Elisabetta Rossi" is "The author of *Happiness*".

When Gianni says «Elisabetta Rossi was the author of *Happiness*», what he says is

- True
- False

Compared to Experiment 1, the alternatives do not include the option "I don't know" because Experiment 2 involves also Gödel cases and not only Jonah ones. The inclusion of that option would have overly complicated the possible combinations of the participants' answers across the two kinds of scenarios.

Let us now examine the specific wording of the vignettes. As a way of example, my analysis will revolve around the Mathematician Case, but analogous considerations also extend to the Painter Case. The first part of the vignette states:

> Heumerus was an ancient Greek mathematician who discovered an important theorem, known as the "prime number theorem". However, Heumerus died shortly after making that discovery, without being able to make it known. Another mathematician, Theodorus, stole the manuscript containing the proof. Theodorus then presented himself as the discoverer of the prime number theorem.

I use the name "Heumerus" to talk about the original bearer of the name, that is, the name's causal-historical referent. This use of the name is not biased against DT, as DT can admit that, within different linguistic communities, such as the narrator's as opposed to the speaker's, a name can have distinct referents in virtue of different associated definite descriptions. I refrain from using the term "suppose", because, as Li (2022, p. 52) argues, «the word "suppose" […] appears to have a strong fanciful and hypothetical flavor which may render keeping track of the stories more difficult» (p. 52). Moreover, as seen, Devitt (2012b) and Devitt and Porot (2018) maintain that a descriptivist narrator would and could not use the expression "suppose" in the context of the Gödel scenarios – although, as argued (section 2.1.1.3), I have some reservation about that thesis. My vignette simply and directly states that Heumerus is not the real discoverer of the mathematical theorem, without introducing that circumstance through the word "suppose". The vignette continues as follows:

> No one ever found out about the theft. As a result, the statement "Theodorus was the discoverer of the prime number theorem" is what everyone believes today; furthermore, this statement is all that has been handed down over time in relation to the name "Theodorus".

This part of the vignette emphasizes the limited information that people have passed down over time. If the vignette had stated that "everyone today believes that Theodorus

is the discoverer of the prime number theorem", the wording would have been incompatible with DT and thus biased against that theory. According to DT, the present-day speakers' beliefs involving "Theodorus" are not about Theodorus, because the associated description does not pick out that individual. The formulation stating that the claim "Theodorus was the discoverer of the prime number theorem" is everything that people passed down over time is neutral, as that claim has undoubtedly been transmitted from speaker to speaker, regardless of the correct theory of reference. An analogous reasoning applies also to the last sentence. I refrain from asserting that the above claim is *about Theodorus*: according to DT, the claim "Theodorus was the discoverer of the prime number theorem", when one of the vignette's present-day speakers believes and utters it, is not about Theodorus, but Heumerus. For that reason, I simply write that that sentence is what people passed down "in relation to the name 'Theodorus'". The last part of the vignette, before the final prompt, reads:

> Marta is a student who took a history of mathematics course, during which she also studied the prime number theorem. Like everyone else, the only description that Marta associates with the name "Theodorus" is "The discoverer of the prime number theorem".

In this final part, I introduce the speaker in semantic contradiction, namely Marta, and further specify the only description that she associates with the name. I also point out that Marta *studied* the incompleteness theorem, to address the possible objection that, within Marta's mental lexicon, the description "The discoverer of the prime number theorem" can pick out some referent only if Marta is able to identify the theorem in the first place, for instance by providing an accurate statement of it.[116] The final prompt is as follows:

> When Marta says «Theodorus was the discoverer of the prime number theorem», what she says is
> - True
> - False

[116] For example, also MMNS's (2004) Gödel Case specifies that «John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem» (p. B6).

The statement that the participant needs to judge is "Theodorus was the discoverer of the prime number theorem", namely, a sentence that juxtaposes the name with its purported reference-fixing description. As argued in section 2.2.2.1, this statement, as opposed for instance to "Theodorus was a mathematical genius", has the advantage of sidestepping a possible attributive interpretation of the name, thereby enabling a more straightforward comparison of DT and CHT. Moreover, the predicate "Being the author of the prime number theorem" is arguably free of any social interpretation (section 2.2.6, n. 87). The sentence is true according to DT and false according to CHT. If a participant engages with this TVJ task by adopting the community's perspective, the sentence is true, since what everyone in Marta's community believes is that "Theodorus was the discoverer of the prime number theorem". Therefore, as argued, in such a context the TVJ "False" can only express a genuinely causal-historical commitment.

3.3.2.2 Jonah-style scenarios

Participants were also presented with two Jonah-style vignettes.

> **Warrior Case**
>
> Ambiorix was one of the administrators of a Gallic tribe, the Eburones. Some descendants of Ambiorix, in order to glorify the memory of their ancestor, spread a story telling that Ambiorix was the warrior who led the Eburones against the Romans when the Romans attacked Gaul. However, in reality, Ambiorix had no warrior training and he spent his entire life in times of peace before the Roman attack. Furthermore, there was no warrior who led the Eburones against the Romans when the Romans attacked Gaul, as the Roman attack caught the Eburones by surprise. No one ever discovered that the story spread among the people was invented. As a result, the statement "Ambiorix was the warrior who led the Eburones against the Romans when the Romans

attacked Gaul" is what everyone believes today; furthermore, this statement is all that has been passed down over time in relation to the name "Ambiorix".

Mattia is a student who took a course on the history of Roman invasions, during which he also studied the Eburones. Like everyone else, the only description that Mattia associates with the name "Ambiorix" is "The warrior who led the Eburones against the Romans when the Romans attacked Gaul".

When Mattia says «Ambiorix existed», what he says is
- True
- False

Astronomer Case

Ugo Vinnoti was an important astronomer of the 16th century who lived in Rome. Two centuries after Ugo Vinnoti's death, during a celebratory event, a biographer gave a presentation about Ugo Vinnoti's life. The biographer narrated that Ugo Vinnoti was used to giving astronomy lessons and that his favorite student was Marco Bianchi. However, in reality, Ugo Vinnoti didn't conduct any lessons and didn't have any students; Ugo Vinnoti had a very individualistic character. Furthermore, Marco Bianchi had no interest in astronomy and wasn't a student of any scientist; he was a simple farmer who neither knew Ugo Vinnoti nor ever lived in Rome. The biographer stated that Marco Bianchi was Ugo Vinnoti's student because the Bianchi family paid the biographer to say so. The Bianchi family, desiring to elevate their status among the families of the time, wanted to make it appear that one of their ancestors had been Ugo Vinnoti's favorite student. That part of the biographer's presentation spread among the people, and no one ever discovered that it was an invented part. As a result, the statement "Marco Bianchi was Ugo Vinnoti's favorite student" is what everyone believes today; furthermore, this statement is all that has been passed down over time in relation to the name "Marco Bianchi".

Leonardo is a student who took a history of astronomy course, during which he also studied the scientific achievements of Ugo Vinnoti. Like everyone

> else, the only description that Leonardo associates with the name "Marco Bianchi" is "Hugo Vinnoti's favorite student".
>
> When Leonardo says «Marco Bianchi existed», what he says is
> - True
> - False

As expected, the Jonah-style vignettes are longer than the Gödel-style counterparts. However, they are shorter than, for example, the Jonah-style scenarios in MMNS (2004) or Islam and Baggio (2020). The minor length should facilitate a smoother understating of the vignettes and avoid the risk that the «probes […] [are] simply too long and complex to generate interpretable data» (MMNS, 2004, p. B7). Let us now proceed by unpacking the vignettes similarly to how I approached the Gödel scenarios. Once again, I will consider a single vignette as a way of example, the Warrior Case, but analogous points extend also to the other vignette, namely the Astronomer Case. Reported below is the first segment of the vignette:

> Ambiorix was one of the administrators of a Gallic tribe, the Eburones. Some descendants of Ambiorix, in order to glorify the memory of their ancestor, spread a story telling that Ambiorix was the warrior who led the Eburones against the Romans when the Romans attacked Gaul. However, in reality, Ambiorix had no warrior training and he spent his entire life in times of peace before the Roman attack. Furthermore, there was no warrior who led the Eburones against the Romans when the Romans attacked Gaul, as the Roman attack caught the Eburones by surprise.

The first part of the vignette specifies that some people spun out certain tales about an individual. One may wonder whether the formulation that "some descendants of Ambiorix […] spread a story according to which Ambiorix was the warrior who conducted the Eburones against the Romans" is consistent with DT, as this wording presents the story as *about Ambiorix*. Theoretically speaking, the situation gets somewhat complicated and subtle at this stage. However, under a few overall reasonable assumptions, this formulation is arguably compatible with DT. The vignette's wording

requires the descriptivist participant to assume that Ambiorix's descendants associate the name "Ambiorix" with some description picking out Ambiorix, thereby making that individual the name's referent. The descriptivist participant can accommodate this requirement by imagining that the descendants know who Ambiorix really was and thus possess a description such as: "Our ancestor who was a peacetime administrator and who actually did this and that instead of being a warrior conducting the Eburones". In such a circumstance, the descendants' uses of "Ambiorix" pick out Ambiorix, since that man is the satisfier of the above description. Therefore, the descendants' utterances are about Ambiorix, to whom they ascribe some invented actions. Moreover, the fabricated content plausibly remains about Ambiorix even among the people who learn the story directly from Ambiorix's descendants. It is possible to imagine that the descendants narrate the invented story by specifying that the individual at the center of it is an ancestor of theirs. Therefore, the addressees come to associate the name not only with the fabricated content, which is true of no individual, but also with some description such as "The Eburonean ancestor whose memory these descendants want to preserve", which is true of Ambiorix. If that is the case, the initial group of recipients who learn the story still understand it as about Ambiorix, since that individual remains the best satisfier of the comprehensive information that those addresses associate with the name. In light of these reasons, *at this stage of the vignette*, the claim "some descendants of Ambiorix […] spread a story according to which Ambiorix was the warrior who conducted the Eburones against the Romans" is arguably not inconsistent with or biased against DT.[117]

[117] On another line, one may object that Ambiorix's descendants' assertions cannot be about Ambiorix because the name "Ambiorix" lacks a pre-existing and widely shared reference-fixing description that links the name to the peacetime leader. As said, the vignette suggests that, before the descendants' narration gets spread, virtually no one knows who Ambiorix is or possesses any lexical item for the name "Ambiorix". That is not surprising, as Ambiorix was a simple administrator with no great achievements on his record. One may argue that, absent this widely shared reference-fixing description picking out the real peacetime administrator, the descendants' use of the name "Ambiorix" cannot be about Ambiorix, since their use does not rely upon any established convention. This criticism unveils one of the most delicate, and in my opinion vulnerable, points of DT: the treatment of the reference-fixing descriptions for names of non-famous people. While a descriptivist can arguably contend that, for example, the widely shared and conventional reference-fixing description for "Giorgia Meloni" is "The current Italian Prime Minister", the purported reference-fixing descriptions for names of non-famous people are way more mysterious. Consider for example my name, "Nicolò D'Agruma". Given that I am not famous, there is no widely shared and established reference-fixing description associated with my name. The people who read my published articles will associate my name with a description like "The author of article X". My parents' friends will associate my name with a description such as "Leonardo and Rosaria's second son". My long-standing friends will associate my name with some conglomerate of descriptions including rich information deriving from our extensive shared experiences. Finally, most people – the extremely vast majority of Italians and habitants of the world – simply do not possess my name in their mental lexicon. The only way DT can accommodate that the uses of my name are not empty is by admitting that, in cases of non-famous people, a description can fix the name's reference even if that description is not widely spread, and instead only a relatively small group of people shares it. Some speakers will refer to me *as* the author of a specific article, some other people *as* the son of a specific couple, and so on. If this reasoning is correct, then DT can and should admit that

Clearly, over the course of time, the situation has changed. The direct contact with Ambiorix's descendants has gradually faded, and what has spread among the people and through successive generations are only the invented deeds. At that point, the story surely is not about Ambiorix anymore, as speakers conventionally associate the name solely with an empty description, that is "The warrior who conducted the Eburones against the Romans". The vignette not only specifies that this description is false of the individual originally called "Ambiorix", but also clearly states that it is empty, as "there was no warrior who conducted the Eburones against the Romans when the Romans attacked the Gauls".[118] Accordingly, as in the Gödel scenarios, I avoid statements to the effect that the present-day people's beliefs are about the original bearer of the name:

> As a result, the statement "Ambiorix was the warrior who led the Eburones against the Romans when the Romans attacked Gaul" is what everyone believes today; furthermore, this statement is all that has been passed down over time in relation to the name "Ambiorix".

Some aspects are worth pointing out. The vignette specifies that the claim "Ambiorix conducted the Eburones against the Romans when the Romans attacked the Gauls" is what everyone today *believes*. One can believe this sentence only if one believes that there existed someone who conducted the Eburones against the Romans. For example, one cannot believe that Louis XVIII was the King of France if one does not believe that there existed someone who was the King of France in the first place. Therefore, that piece of the vignette entails that people believe that "Ambiorix existed" and thus this sentence is one that they can assert, as it in fact happens in the final prompt. Moreover, the description that all the speakers associate with the name does not outline any prototypically fictional entity or any character whose actions lie beyond the practical abilities of real individuals. Instead, the description outlines a warrior who conducted an army against another force: warlords are customary figures within battlefield contexts.

---

the descendants' uses of "Ambiorix" refer to Ambiorix, as that man is the satisfier of the description that a specific subset of people, namely Ambiorix's descendants, associate with the name.

[118] As regards this point, my vignette differs from Devitt and Porot's (2018), as theirs states that the tales that speakers associate with "Ambiorix" are false of Ambiorix, but does not explicitly specify that they are false of *any* individual. Devitt and Porot assert that those tales are legends, which suggests that no person satisfies them. Likely, their failure to state that no individual satisfies the tales had no bearing upon the results that they collect. That said, the addition of this specification into my wording makes the contrasting predictions deriving from DT and CHT clearer.

This portrayal refrains from introducing extravagant elements, particularly when eschewing statements suggesting that the warlord engaged in ostensibly superhuman feats, such as simultaneously slaying numerous adversaries or similar actions. In this way, it is even clearer that anyone can and do believe that "Ambiorix existed". Furthermore, the fact that the description avoids outlining a prototypically fictional entity helps circumvent the problem of fictional incorporation (section 1.5), thereby clearly establishing the causal-historical prediction that "Ambiorix existed" is true. Finally, the vignette specifies that the claim "Ambiorix was the warrior who conducted the Eburones" is the *only* piece of information that current speakers associate with the name.[119] As a consequence, DT clearly predicts that the name is empty: no other possible piece of information might link the name to some real entity.

The final prompt reads as follows:

> Mattia is a student who took a course on the history of Roman invasions, during which he also studied the Eburones. Like everyone else, the only description that Mattia associates with the name "Ambiorix" is "The warrior who led the Eburones against the Romans when they attacked Gaul".
>
> When Mattia says «Ambiorix existed», what he says is
> - True
> - False

The final prompt introduces the speaker in semantic contradiction, Mattia. Furthermore, it explicitly specifies that Mattia derives his information from a course centered around the Roman invasions. This additional emphasis further stresses that Mattia truly believes that "Ambiorix was the warrior who conducted the Eburones" and thus that "Ambiorix existed", as a scholastic or university course is a source that people regard as reliable for belief formation. The sentence that Mattia utters is "Ambiorix existed", which is true according to CHT, because there existed an individual at the end of the name's causal-historical chain, but is false according to DT, because there was no individual satisfying the description that Mattia associates with the name. If a participant interprets the truth

[119] Machery (2021, p. 546) too stresses the relevance of specifying that a certain description (e.g., "The discoverer of the incompleteness theorem") is the *only one* that speakers associate with the critical name (e.g., "Gödel").

predicate from the community's perspective, she will answer that the sentence is true, because the consensus within Mattia's community is that "Ambiorix existed" and that individual was the warrior who conducted the Eburones. Therefore, as argued, in such a context the TVJ "False" can only express a genuinely descriptivist commitment.

3.3.2.3 Fillers

Experiment 2 includes two fillers. Analogously to Experiment 1, the fillers unambiguously admit only one answer as correct and also function as comprehension and attention checks, as they control that the subject pays attention to the sequence of vignettes over the course of the experiment. Compared to Experiment 1, two differences emerge. First, I included only one question into the filler, aiming to enhance simplicity and reduce the cognitive burden on the participants. Second, in line with the options available for the targets, the filler admits only "True" and "False" as answers (that is, the "I don't know" option is not available). Reproduced below are the two fillers.

Football Player filler

> Antonio Salvi was an Italian footballer from the 1960s who played in the Argentine football championship. Antonio Salvi passed away in the 1990s. He was a versatile forward, strong in heading and capable of using both feet to kick. Antonio Salvi won the top scorer title in the Argentine championship several times. After his football career, Antonio Salvi also pursued a career as a coach. However, as a coach, Antonio Salvi didn't achieve as dazzling a success as he did when he was a player on the field. A very important aspect of Antonio Salvi's life that made him famous was his philanthropic activity: Antonio Salvi used to donate a portion of the money he earned, both as a

player and later as a coach. The causes he championed the most were the construction of hospitals and schools for the underprivileged classes.

The story of Antonio Salvi is often told to young people nowadays as an example of the good that can be done with one's earnings.

"The success of Antonio Salvi as a coach was inferior to the success of Antonio Salvi as a footballer". Based on what you have read, this statement is

- True

- False

Poet filler

Chiara Pedrini was an Italian poet of the 20th century. Her poems were highly appreciated, and her poetry collections were consistently top sellers. For many years, Miuka was Chiara Pedrini's companion dog, and some of her poems were inspired by Miuka. Chiara Pedrini was deeply fond of Miuka, to the extent that when the animal passed away, she decided not to have any other pets. In fact, Chiara Pedrini believed that no one could replace Miuka. Following that loss, Chiara Pedrini's artistic inspiration dwindled. Many of her new poetry collections received less acclaim, and she lost much of her fame as a successful poetess.

The story of Chiara Pedrini is often told to young people nowadays as an example of how humans can be fond of animals.

"Miuka was Chiara Pedrini's companion cat for many years". Based on what you have read, this statement is

- True

- False

The only answer that the Football Player filler admits as correct is "True", while the only answer that the Poet filler admits as correct is "False".[120]

3.3.3 Procedure

I administered the experiment online, using PsyToolKit (Stoet, 2010, 2017). Following an introductory screen, participants are asked to provide their consent to participate in the experiment. At this point, they receive an attention check.[121] Following that, they are presented with the four targets and the two fillers in random order. The TVJ final prompt at the end of the target or fillers is on the same page as the item itself. After that, participants answer a demographic questionnaire. Finally, the participant needs to further confirm that she wants to send her answer. Absent that additional consent, the system automatically excludes the answer.

[120] In her comment at the end of the questionnaire, one participant noted that, contrary to my assessment, the final sentence of the Football Player filler, that is "The success of Antonio Salvi as a coach was inferior to the success of Antonio Salvi as a footballer", might admit "False" as the correct TVJ. Although the filler clearly asserts that "as a coach, Antonio Salvi didn't achieve as dazzling a success as he did when he was a player on the field", the participant observed that personal success extends beyond professional achievements (in this case, in the field of sport), but can also encompass a moral dimension, involving the good that one contributes to society. Given that Antonio Salvi gave part of his earnings that he obtained "both as a player and later as a coach", he achieved great moral success in both roles. As a result, the participant noted that there is no reason to claim that Antonio Salvi's success as a coach was inferior to his success as a player, and thus she answered "False". While developing the filler, I did not consider this possible reading of the final prompt. 9 participants opted for the TVJ "False" in the context of the Football Player filler. Out of those, 3 would not have been included in the final analysis anyway – e.g., for failing the attention check that appears at the beginning of the entire survey. I did not include the remaining 6 participants in the analyses. However, their inclusion exerts virtually no influence on the resulting proportions and statistics, as I also show in the results section (3.3.5).

[121] The attention check reads as follows:

> Below is a question about your literary tastes. In reality, this question is a test to evaluate whether the participant pays attention to what he or she read. Regardless of what your literary preferences are, I kindly ask you to respond by selecting the "Other" option and filling in the box with the phrase "I have read the instructions" ["Ho letto le istruzioni"] (with double quotation marks and the capital "H": for your convenience, I recommend that you copy and paste).
>
> Given the instructions above, I kindly ask you to answer the following question: what is your favorite literary genre?
>
> - Biography
> - Science fiction
> - Crime
> - Poetry
> - Essay
> - Other

3.3.4 Participants

249 Italian participants were recruited through Prolific. From this initial pool, the sample included in the final analysis meets the following criteria.

(a) They are native Italian speakers.[122]
(b) They answered the initial attention check correctly.
(c) They answered both fillers correctly.
(d) They do not have a university degree in Philosophy.
(e) They did not attend a course in Philosophy of Language.

Among the participants involved in the study, 46 participants (18.47%) do not meet one or more of the inclusion criteria. More specifically, 21 (8.43%) participants did not pass the attention check. 15 participants (6.02%) did not answer one or both the filler tasks correctly. 3 participants (1.20%) have a university degree in Philosophy and 15 (6.02%) attended a course in Philosophy of Language. 203 participants are included in the final analysis ($M = 30.75$; $SD = 8.93$; 106 females, 95 males, 2 prefer not to specify) and took an average of 8.17 minutes to complete the test, with a median time of 7 minutes.

3.3.5 Results

As regards the two Gödel cases, the Mathematician Case and the Painter Case, 85.22% and 87.68% of participants choose the TVJ "False", respectively. The Chi-square statistics confirm that the difference in frequency for each vignette is statistically significant (respectively: χ^2 (1, N = 203) = 100.73, $p < .001$; χ^2 (1, N = 203) = 115.32, $p < .001$). The McNemar test confirms the alignment between the two cases, since no statistically significant difference emerges ($p > .05$). By aggregating the two vignettes, participants express the TVJ "False" 86.45% of the time, while the TVJ "True" 13.55%

[122] I adopted pre-screener filters on Prolific such that participants: are Italian, are born in Italy, are native Italian speakers and spent most of their life in Italy before turning 18. In the final demographic questionnaire, I re-validated the Italian speaker condition by asking participants "Are you a native Italian speaker?", and I also asked them "Did you spend most of your life in Italy?". All participants answered "Yes" to both questions.

of the time (Figure 3.3). The Chi-square statistics confirm that the difference in frequency is significant (χ^2 (1, N = 406) = 215.80, $p < .001$).

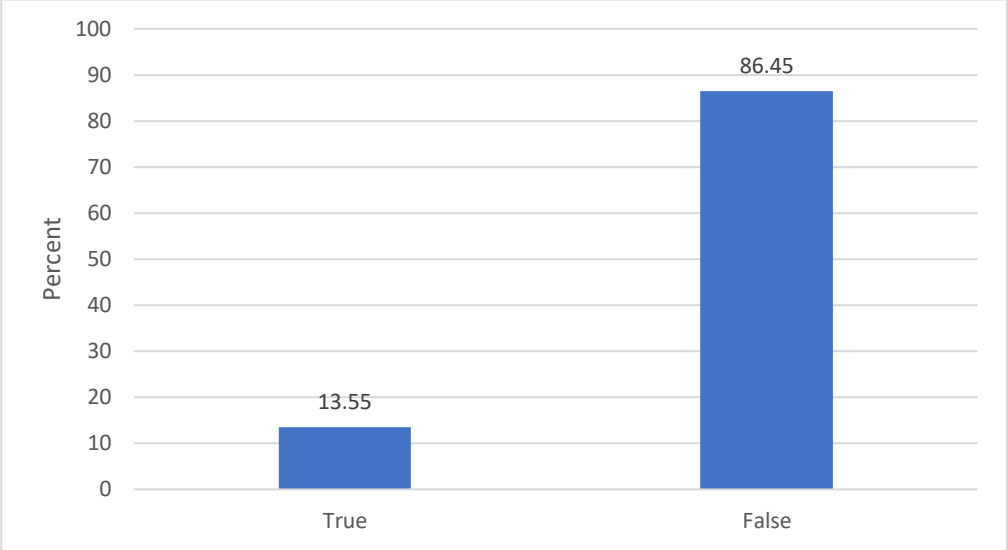

Figure 3.3 Aggregated TVJs of the two Gödel cases in Experiment 2

As regards the two Jonah cases, the Warrior Case and the Astronomer Case, 91.63% and 90.15% of participants choose the TVJ "True", respectively. The Chi-square statistics confirm that the difference in frequency for each vignette is statistically significant (respectively: χ^2 (1, N = 203) = 140.70, $p < .001$; χ^2 (1, N = 203) = 130.88, $p < .001$). The McNemar test confirms the alignment between the two cases, since no statistically significant difference emerges ($p > .05$). By aggregating the two vignettes, participants express the TVJ "True" 90.89% of the time, while the TVJ "False" 9.11% of the time

(Figure 3.4). The Chi-square statistics confirm that the difference in frequency is significant (χ^2 (1, N = 406) = 271.49, $p < .001$).

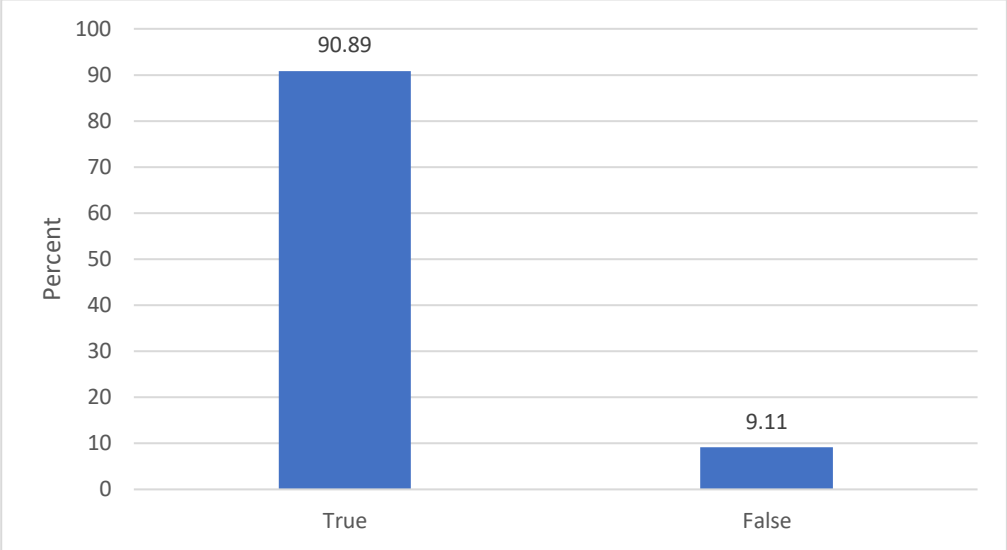

Figure 3.4 Aggregated TVJs of the two Jonah cases in Experiment 2

Overall, the kinds of scenarios (Gödel vs Jonah) elicit clearly different TVJs, as the two Gödel vignettes mainly prompt the TVJ "False", while the two Jonah vignettes prompt the TVJ "True". That is also confirmed by a Cochran's Q test, which indicates a statistically significant difference across the four vignettes ($p < .001$). Given the similarity between the two Gödel Cases and between the two Jonah ones, confirmed by the lack of statistical significance of the respective McNemar tests, what drives the statistical significance of the Cochran's Q Test is the difference between each Gödel vignette with each Jonah vignette. Indeed, McNemar tests show that the Mathematician Case differs significantly from both the Warrior Case and the Astronomer Case, and that the Painter Case differs significantly from the Warrior Case and the Astronomer Case one (all *p*s < .001).

As discussed, presenting both Gödel vignettes and Jonah ones to the participants allow us to directly observe the participants' patterns of answers across the two kinds of scenarios. I divided the participants' answers into nine categories, depending on the pattern that they express. Within the context of the Gödel scenarios, a participant can express three different combinations: she can choose the TVJ "True" for both the Gödel scenarios, in which case I label the participant's answer as "True"; she can choose one

TVJ “True” and one TVJ “False”, in which case I label her answer as “Mixed”; she can express the TVJ “False” in both scenarios, in which case I label her answer as “False”. The same classification extends to the Jonah cases. Combining the participant’s pattern of answers in the Gödel Case with the one in the Jonah Case provides the participant’s overall resulting pattern. A pattern “G: True / J: True” signals that the participant always expressed the TVJ “True” across the four vignettes. A pattern “G: Mixed / J: True” signals that she expressed one TVJ “True” and one “False” in the context of the Gödel Case, while she expressed the TVJ “True” for both the Jonah scenarios. Table 3.7 reports the nine possible combinations:

Table 3.7 Possible patterns of answers across the Gödel and Jonah Case in Experiment 2

Pattern	**Gödel Case**	**Jonah Case**
A	False	True
B	True	True
C	Mixed	True
D	True	Mixed
E	Mixed	False
F	Mixed	Mixed
G	False	False
H	False	Mixed
I	True	False

Figure 3.5 summarizes the resulting patterns:

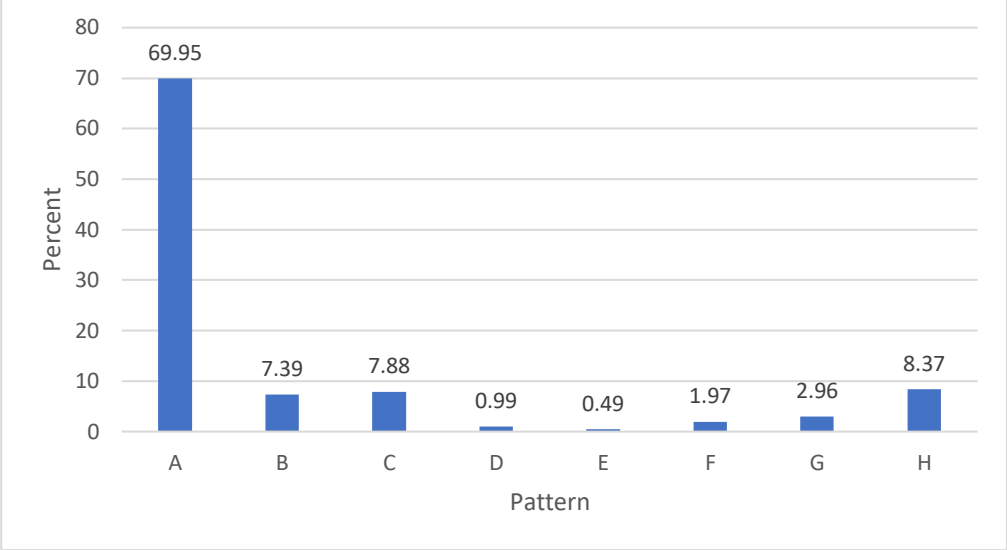

Figure 3.5 Patterns of answers in Experiment 2

69.95% of participants express the combination A, 7.39% the combination B, 7.88% the combination C, 0.99% the combination D, 0.49% the combination E, 1.97% the combination F, 2.96% the combination G and 8.37% the combination H.[123] No participant opts for pattern I.[124]

[123] The Chi-square statistics confirm the frequency of A differs from all the other options: A vs. B (χ^2 (1, N = 157) = 102.73, $p < .001$); A vs. C (χ^2 (1, N = 158) = 100.48, $p < .001$); A vs. D (χ^2 (1, N = 144) = 136.11, $p < .001$); A vs. E (χ^2 (1, N = 143) = 139.03, $p < .001$); A vs. F (χ^2 (1, N = 146) = 130.44, $p < .001$); A vs. G (χ^2 (1, N = 148) = 124.97, $p < .001$); A vs. H (χ^2 (1, N = 159) = 98.27, $p < .001$). The Chi-square statistics confirm the frequency of B differs from D (χ^2 (1, N = 17) = 9.94, $p < .01$), from E (χ^2 (1, N = 16) = 12.25, $p < .001$), from F (χ^2 (1, N = 19) = 6.37, $p < .05$) and from G (χ^2 (1, N = 21) = 3.86, $p < .05$). The frequency of C differs from D (χ^2 (1, N = 18) = 10.89, $p < .001$), from E (χ^2 (1, N = 17) = 13.24, $p < .001$), from F (χ^2 (1, N = 20) = 7.20, $p < .01$) and from G (χ^2 (1, N = 22) = 4.55, $p < .05$). The frequency of H differs from D (χ^2 (1, N = 19) = 11.84, $p < .001$), E (χ^2 (1, N = 18) = 14.22, $p < .001$), F (χ^2 (1, N = 21) = 8.05, $p < .01$) and G (χ^2 (1, N = 23) = 5.26 $p < .05$). The comparisons between D, E, F and G were not run due to the limited number of participants falling under these patterns, respectively 2, 1, 4 and 6. All other comparisons are not significant (all *p*s > 0.05).

[124] As previously mentioned, there might be reason to include the 6 participants whom I have excluded because of their answer to the Football Player filler. Their inclusion has an inconsequential impact. Including these participants, the sample size increases to 209 (M = 30.61; SD = 8.85; 109 females, 98 males, 2 prefer not to specify) and they took an average of 8.13 minutes to complete the test, with a median time of 7 minutes. As regards the Gödel scenarios, the Mathematician Case elicits 14.35% of TVJs "True" and 85.65% "False", and the Painter Case elicits 12.44% of TVJs "True" and 87.56% "False". As regards the Jonah vignettes, the Astronomer Case elicits 89.47% of TVJs "True" and 10.53% "False", and the Warrior Case elicits 91.87% of TVJs "True" and 8.13% "False". Overall, the Gödel cases elicit 13.40% of TVJs "true" and 86.60% "False", while the Jonah cases 90.67% of TVJs "True" and 9.33% "False". The pattern distribution is the following: A: 69.86%; B: 7.18%; C: 7.66%; D: 0.96%; E: 0.48%; F: 2.39%; G: 2.87; H: 8.61%; I: 0%. All the statistical significances and non-significances, including the levels of significance, remain unaffected for all the analyses, except for the comparison between C and F, which shifted from being significant at the .01 level to the .05 level.
For the sake of completeness, this note details the analyses for the entire group of participants, including also those who do not meet the inclusion criteria. The general picture of the results overall mirrors the one deriving from the final sample. 249 participants completed the survey (M = 30.55; SD = 8.72; 127 females, 120 males, 2 prefer not to specify)

3.3.6 Discussion

The collected results support CHT as the correct theory of reference. To begin, let us examine the distributions of the aggregated TVJs within each kind of scenario – the Gödel Case and the Jonah Case. The vast majority of participants express the TVJs predicted by CHT: the TVJ "False" in the context of the Gödel Case and the TVJ "True" in the context of the Jonah Case. As regards specifically the Gödel Case, the data are in line with those that Li (2021) collect, where 90% of her Western sample (Americans) expresses the TVJ that CHT predicts. This result corroborates the thesis that the proportion of causal-historical TVJs in the Gödel Case is higher in response to vignettes avoiding the confounds of the social construal of predicates and the attributive interpretation of names.

Let us now analyze the patterns. A prevails (69.95%) and differs statistically significantly from all the other patterns. The other three patterns that participants express more frequently are H (8.37), C (7.88) and B (7.39), which do not differ significantly among them but they do from all the other patterns. All the remaining patterns collect very few participants, ranging from 0.49% to 2.96% of the total sample. The distribution of the participants' patterns corroborates CHT as the correct theory of reference for proper names. The highly most expressed pattern, namely A, is clearly causal-historical: the participant chooses the TVJ "False" for both the Gödel scenarios and the TVJ "True" for both the Jonah scenarios, which is precisely what one would expect from a Kripkean participant judging the factual truth value of the sentences.

As said, after A, the most chosen patterns are B, C and H. Two of them, B and C, are the patterns that one would expect from a causal-historical participant who adopts the community's perspective in the Gödel Case (for one or both the vignettes). These participants choose the TVJ "True" for one or both the Gödel scenarios and the TVJ "True" for both the Jonah scenarios. This finding is in line with Hypothesis 2, namely the

---

and took an average of 8.25 minutes to complete the test, with a median time of 7 minutes. As regards the Gödel scenarios, the Mathematician Case elicits 14.46% of TVJs "True" and 85.54% "False", and the Painter Case elicits 12.05% of TVJs "True" and 87.95% "False". As regards the Jonah vignettes, the Astronomer Case elicits 89.16% of TVJs "True" and 10.84% "False", and the Warrior Case 90.36% of TVJs "True" and 9.64% "False". Overall, the Gödel cases elicit 13.25% of TVJs "True" and 86.75% "False", while the Jonah cases 89.76% of TVJs "True" and 10.24% "False". The pattern distribution is this: A: 69.08%; B: 6.43%; C: 8.43%; D: 1.20%; E: 0.80%; F: 2.01%; G: 3.61; H: 8.43%; I: 0%. All the statistical significances and non-significances, including the levels of significance, remain unaffected for all the analyses, except for the comparison between B and G, which is no longer significant. Thus, the results obtained from the entire group of survey respondents are broadly consistent with those from the sample included in the final analysis.

hypothesis that the participants choosing the TVJ "True" in the Gödel Case do not express a genuine descriptivist answer. If those participants were providing their TVJs out of a genuine descriptivist commitment, they should choose the TVJ "False" in the context of the Jonah Case, as in the Jonah Case the TVJ "False" can only voice a descriptivist answer. More in general, the results show that most of the residual amount of TVJs "True" stemming from the Gödel Case, even after controlling for the potential social interpretation of the predicate and the attributive understanding of the name, is due to the participants' adoption of the community's perspective. Among the entire sample, 18.72% of participants express the TVJ "True" across the two Gödel scenarios at least once. However, the vast majority of those participants, 81.58%, choose only the TVJ "True" across the two Jonah scenarios, thereby fitting within patterns B and C, as expected if their TVJs in the Gödel Case do not express a genuine descriptivist answer. The residual 18.42% of that sub-sample falls under the fragmented options D (0.99%), E (0.49%) and F (1.97%).

As mentioned, another pattern, comparable in prevalence to B (7.39%) and C (7.88%), is H (8.37%), in which participants choose the TVJ "False" for both the Gödel cases and express mixed answers in the context of the two Jonah cases. The causal-historical monist hypothesis cannot account for pattern H, since a Kripkean theorist cannot accommodate any TVJ "False" in the Jonah Case. As argued, from a Kripkean point of view, in the context of the Jonah Case the widely shared consensus reflects the factual truth, both pointing to the TVJ "True". Therefore, a causal-historical participant has no reason to divert from the TVJ "True" that CHT predicts and rather choose the TVJ "False". It is probable that the responses from the participants falling under H reflect a certain degree of experimental noise. The critical sentences of the Jonah Case, such as "Ambiorix existed", are arguably less natural and ordinary than their counterparts in the Gödel Case. If one needs to explain who Aristotle was to a friend who never heard of him, one will not claim – for instance – that "Aristotle existed and was Plato's most brilliant student". Rather, one will simply claim that "Aristotle was Plato's most brilliant student", presupposing that one is talking about a real human being. As argued (section 3.1.2), the choice of using a sentence like "Ambiorix existed" stems from the clearly divergent predictions that DT and CHT make about this sentence's truth value and in light of the predicaments concerning other possible critical sentences. However, this methodological

desideratum comes with a trade-off in terms of naturalness. The use of a somewhat artificial sentence may have led some participants to approach the judgment task with a degree of perplexity or disorientation, thereby introducing some level of noise into their responses.

One can object that there are alternative hypotheses consistent with the data, such as a pluralist position according to which names are ambiguous between two referential mechanisms.[125] This interpretation can accommodate the data by stipulating that 70% of participants – those expressing pattern A – consistently adopt the causal-historical referential mechanism, while the remaining 30% –falling within all the other patterns, especially B, C and H – shift from one referential mechanism to the other. Naturally, a pluralist position can accommodate *any* kind of variation. Given a fluctuation between *a* and *b* (where *a* and *b* can be intuitions, TVJs or other measurable indices), the pluralist can easily assert that two theories, X and Y, are both valid and *a* aligns with the predictions of X while *b* aligns with the predictions of Y. In light of that, the crucial question is not whether my results are consistent with a pluralist position, but rather whether there are compelling grounds to favor the pluralist position over a monist one.[126] A pluralist position offers a less convincing explanation of the data when contrasted to the monist Kripkean position including the possibility to adopt the community's perspective. If names really had two reference-fixing mechanisms, then one would expect not only some participants voicing a purely causal-historical pattern (A), but also some other participants expressing a purely descriptivist pattern (I). However, it is noteworthy that not a single subject manifests pattern I, namely chooses "True" in both the Gödel scenarios and "False" in both the Jonah scenarios. As MMNS (2004) suggest, cultural influences might lead participants from a Western cultural background, such as our Italian sample, to favor the causal-historical understanding of names over the descriptivist one. However, the circumstance that the vast majority of the sample – 70% – consistently opts for only one of the two referential mechanisms, while literally no participant exhibits the opposite behavior, is a finding that the pluralist theorist can difficulty explain.

[125] Or a pluralist, hybrid position, according to which a proper name has only one referential mechanism that includes both a descriptivist element and a causal-historical one (section 2.3.2.3). My critical considerations apply to a hybrid theory as well.

[126] Kripke claims that «it is very much the lazy man's approach in philosophy to posit ambiguities when in trouble» (p. 268). As I am about to argue, the issue at hand is not necessarily one of laziness, but rather a matter of overall fit of the pluralist position with the collected data.

In conclusion, Experiment 2 corroborates two theses. First, a good portion of the supposedly descriptivist evidence that Italian samples[127] provide in the context of the Gödel Case stems from confounding factors like the attributive understanding of names or the social interpretation of predicates. Second, the residual variation among Italian participants as regards the Gödel Case does not have a semantic root; specifically, the variation stems from a perspectival ambiguity and not from a shift between descriptivist and causal-historical name usage. Therefore, Experiment 2 validates the results of Experiment 1, with an improved methodology. Moreover, although Experiment 2 does not gather cross-cultural data, the experiment has an indirect cross-cultural significance. If an experiment, like Experiment 2, shows that the intra-cultural variation within one group lacks a semantic basis, that very same work questions the supposed semantic nature of the variation within and across other samples. Results like those of Experiment 2 further emphasize that, regardless of the group under investigation, any experiment has to control for the impact of the epistemic ambiguity. Otherwise, it is naïve or simplistic to attribute a semantic relevance to any intra or cross-cultural variation.

[127] Vignolo and Domaneschi (2018, 2022) and Domaneschi and Vignolo (2020).

4 Experiment 3: an eye-tracking study on the Gödel Case

The most effective way to test theories of reference is to collect data about linguistic usage, which comprises production and comprehension. In the context of a Gödel Case, TVJs can capture linguistic comprehension because they shed light on the proposition that a participant understands as expressed by John's claim. However, TVJs are epistemically ambiguous. Scholars can try to control for that confounding aspect by asking follow-up questions; however, that strategy may render the final prompt too theoretical, thereby rekindling Martí's (2009) criticism. In Experiment 2, I have tried to control for the epistemic ambiguity by avoiding follow-up questions and rather combining the TVJs on the Gödel Case with those on the Jonah Case, following Islam and Baggio's (2020) strategy.[128] The results suggest that, among Italian participants and potentially extending to Westerners more broadly, CHT is the correct theory of reference and the residual purportedly descriptivist TVJs that the Gödel Case elicits stem from the participants' adoption of the speaker's perspective – or, more precisely, of the perspective of the speaker's community. However, TVJs have faced another criticism, according to which a TVJ on a sentence uttered by a third speaker, even when devoid of any follow-up question, would suffer from precisely the same problem that Martí highlights with regard to RIs. Beyond any potential epistemic ambiguity, a TVJ final prompt would be problematic because it invites the participant to reflect on how a speaker (John, Ivy or whoever else) uses the name, a task that belongs to the semanticist, and hence is theoretical. Therefore, the participant expresses her TVJ on the basis of the view – DT or CHT – that she finds theoretically more intuitive, rather than on her own understanding of John's claim. In the present section, I analyze and evaluate this criticism and present an eye-tracking study, coauthored by Filippo Domaneschi, Massimiliano Vignolo, Camilo Rodríguez Ronderos and myself, that aims to test linguistic comprehension without necessitating reliance upon TVJs.[129]

[128] That is also partially true for Experiment 1, given the ultimately secondary role that Experiment 1 ascribes to the follow-up questions.

[129] As said, the eye-tracking study is a coauthored work. For that reason, while presenting the experiment (from section 4.2 to 4.6), I use the inclusive pronoun "we". However, while presenting the theoretical framework that contextualizes the experiment (section 4.1), I restrict myself to the pronoun "I", because my coauthors not necessarily agree with all the details of my analysis. That said, they share the fundamental considerations integral to the study's foundation.

4.1 Background

4.1.1 Does the presence of a speaker necessarily imply theoretical reflection?

Martí presents a criticism against MOD's (2009) use of TVJs that is conceptually analogous to the one that she raises against the original MMNS (2004). MOD's study too is too theoretical and thus fail to test linguistic usage: «my concerns with the new probe [the one that MOD (2009) propose] are pretty much the same concerns I had with the first study [MMNS (2004)]» (Martí, 2012, p. 74). Rather than choosing the TVJ on the basis of the proposition that she understands as expressed by John, a participant would perform a TVJ final prompt in a different way. She reasons abstractly about the name "Gödel" as John uses it, chooses the theoretical explanation that she finds more intuitive for explaining how the name can pick out an extralinguistic entity, and finally expresses a TVJ on the basis of that theoretical reflection. According to Martí, a participant may engage in that reasoning because the TVJ final prompt still presents the critical sentence as uttered by *someone else*, like John or Ivy.

> The question I proposed in Martí 2009 is meant to elicit responses in which subjects use 'Gödel'; it does not ask subjects to reflect on *someone's use* of 'Gödel'. Does MOD's linguistic question do that? It seems to me that it does not, for […] the questions still ask subjects to reflect on someone's practice using 'Gödel'; it doesn't require of them to use 'Gödel'.
> (Martí, 2012, pp. 74-5, my emphasis)[130]

Presenting to the participant a sentence that someone else utters would involve a reflection on the linguistic usage of that person. Therefore, a test like MOD's would fail to unveil linguistic usage.

Note that the primary source of the problem, according to Martí, is not the use of TVJs, but rather the presence of a third speaker, like John or Ivy. Indeed, Martí claims that TVJ tests like the ones in Devitt and Porot's (2018), which exclude speakers such as John or

[130] For easy reference, I report below the question that Martí proposes at the end of her (2009), which I already quote in section 1.4.3.4:

> One day, the fraud is exposed, and John exclaims: «Today is a sad day: we have found out that Gödel was a thief and a liar».
> What do you think of John's reaction?
> (Martí, 2009, p. 47)

Ivy, are closer to the raw linguistic data against which scholars should test theories of reference.

> Devitt and Porot make some important, and very much needed, changes to the vignette. […] [That] change [is] the elimination of Ivy. In Devitt and Porot's vignette there is no mention of a speaker (such as the Ivy of MOD or the John of MMNS) whose use participants in the experiment are required to think about. […] the elimination of Ivy is an important contributing factor that makes the test closer to actual use. MOD's questions are not as close to actual use as Devitt and Porot suggest, for they still require explicitly that the subjects think about how another speaker is using a name and what she is referring to when she uses it, a reflection that, as I have argued, is theoretical, and hence it is not obvious that it provides direct evidence of disposition to use. […]
> (Martí, 2020, p. 335)

However, Martí adds that, even when devoid of a speaker as in the case of Devitt and Porot (2018), TVJs are «imperfect tests of usage» when compared to linguistic production (Martí, 2020, p. 335). That is because a TVJ prompt asks the participant to interpret a sentence, and that is a process where pragmatic principles like «charity and accommodation often play an important role», thereby potentially preventing the TVJ from properly reflecting the proposition that the sentence semantically encodes (*ivi*, p. 334).

One can raise two objections. First, the elimination of the speaker, like John or Ivy, exposes the TVJ test to the New Meaning Objection, as the description upon which the descriptivist participant is supposed to rely is ambiguous. As discussed in the previous sections of the thesis, the state of semantic contradiction of a speaker like John or Ivy is essential for designing a test that elicits different predictions from DT and CHT. Second, the thesis that principles like those of accommodation or charity can taint only interpretation data and not also production data is debatable. Consider the following example. While it is widely known among biologists that spiders are not insects, the general public often holds the misconception that they are. Therefore, even a biologist, when entertaining a casual conversation with people outside her field, may follow the principle of charity by adhering to the ordinary level of biological knowledge and

applying the term "insect" to spiders as well. Therefore, Martí's sharp demarcation between interpretation and production in terms of the application of cooperative principles may not be fully grounded.

That said, let us come back to Martí's idea that the presence of a speaker, like John or Ivy, entails theoretical reflection on the participant's side. Some authors object to that thesis (Cohnitz, 2015; Cohnitz and Haukioja, 2015; Devitt and Porot, 2018) and argue that the presence of a speaker like John or Ivy does not ordinarily require the participant to develop any explicit theoretical reflection on the words that that speaker uses, proper names included. If John says to Smith "Federer is the best player in tennis history", there seems to be no evidence that Smith needs to engage in conscious theoretical reflections on how John uses the single terms of the above utterance, such as "Federer", "was", "tennis" and so on. Simply, Smith listens to John and processes the proposition that John's utterance conveys, namely, a proposition about who the best tennis player in history is. Cohnitz (2015) stresses the point:

> Linguistic competence with proper names isn't only a matter of producing sentences with proper names in the right way. It's also a matter of interpreting the use of proper names in the sentences of others. Presumably production and interpretation are just two sides of the same coin. Therefore the fact that the question asked in the probe concerned *the utterance of a third person* [such as John or Ivy] does not itself establish that the probes were collecting anything other than the relevant "raw data" [Cohnitz takes the expression "raw data" from Martí (2014, p. 22), to mean "untheoretical data"].
> (Cohnitz, 2015, p. 96, my emphasis)[131]

Devitt and Porot argue for the same thesis, by referencing some works in psycholinguistics:

> This sort of understanding of another's [e.g., John's or Ivy's] language is an ordinary exercise of our linguistic competence. We learn from psycholinguistics that it

[131] Cohnitz develops those considerations in the context of a hypothetical vignette in which John utters the sentence "Gödel is a brilliant mathematician" and the participant has to answer the question "What does 'Gödel' refer to in this utterance?" (p. 95). That setting deviates from the ones that scholars actually employed in the literature, where, when a participant has to express a RI judgment on the name "Gödel", she is not presented with any utterance by John. (The only exceptions are Experiment 1 in Lam (2010), which however Cohnitz does not mention, and MMNS's (2004)

> involves subconscious, subpersonal, automatic, extraordinarily fast processing, and that is mostly all that it involves. (Gernsbacher and Kaschak, 2003; Pickering, 2003; Tanenhaus, 2003 [see also Schlesewsky and Friederici, 2003]).
> (Devitt and Porot, 2018, p. 1560)[132]

Naturally, the thesis that processing what someone else, like John, says does not normally involve conscious theoretical reflection is consistent with the idea that the result of that process, namely the apprehension of the proposition that John expresses, is consciously accessible to the participant or listener:

> In the case of semantic interpretations, i.e. our understanding of utterances and their parts, it seems plausible to assume that the results of these interpretations are available to consciousness. After all, these results, that is, our interpretations of what

Jonah 1 and Jonah 2, which are not Gödel scenarios.) That said, Cohnitz and Haukioja (2015), while discussing Martí's (2009) criticism, explicitly maintain that also a classical RI prompt, like the one in MMNS (2004), can potentially be a test of linguistic usage. Cohnitz and Haukioja's rationale for their claim is that, irrespective of whether a participant is presented with any utterance by John, she can *imagine* John as enunciating a name-containing sentence and ground her RI on the proposition that she understands as expressed by that claim. In that case, performing a RI prompt would be similar to addressing a TVJ one, since in both cases the participant is supposed to formulate her answer on the same basis – the proposition that she associates with John's claim. Cohnitz and Haukioja's reconstruction outlines a way in which a participant can potentially address a RI final prompt. However, as I am also poised to stress, RIs remain arguably more problematic than TVJs, as the RI test explicitly asks the participant whom the name, as John uses it, refers to. Thus, the prompt arguably invites the participant to engage in a theoretical reflection on the name itself rather than in the comprehension of an imagined name-containing claim that John may utter. The situation markedly contrasts with a TVJ final prompt, which asks no explicit question about the name and presents the participant with a sentence that John in fact claims (when John says "Gödel is a great mathematician"…), as opposed to a hypothetical sentence with which the participant is supposed to enrich the vignette.

[132] As said, Devitt and Porot TVJ tests do not include the speaker in semantic contradiction. However, that is a feature that Devitt and Porot do not emphasize, as Martí (2020, p. 335) notes: «there is another change whose importance, in my view, Devitt and Porot fail to appreciate fully: the elimination of Ivy». Precisely because of that lack of emphasis, Devitt and Porot do not consider Ivy's presence as a factor invalidating MOD's (2009) test. Indeed, when Devitt and Porot reflect on why MOD's results differ from their own, they consider that MOD's findings are either anomalous or subject to some wording effects, but they do not explicitly identify Ivy as the source of the difference: «What should we conclude about MOD's findings? It is possible, of course, that the use of proper names among Americans differs from that among Indians, Mongolians, and the French. This needs to be tested but it strikes us as very unlikely. Thus, we conclude that either MOD's findings are anomalous, or else the specific prompt and vignette used in their study were insensitive to a real pattern in the usage of proper names. The latter is compatible with a susceptibility to wording effects for TVJ tests» (Devitt and Porot, 2018, p. 1575). For all these reasons, although Devitt and Porot's test omits speakers like Ivy, they do not seem in principle against her presence. Indeed, the psycholinguistic excerpts that they mention and the arguments that they develop are suitable for all the TVJ tests, also those that incorporate the speaker in semantic contradiction. However, it is worth noting that, in his later (2023), Devitt does maintain that the elimination of the vignette's character improves the design of the experiment (pp. 1153 and 1155). That elimination circumvents the confounding factors that may influence the TVJ prompt, such as the evaluation of whether Ivy's utterance aligns with the consensus within her community. As I am about to stress, the eye-tracking methodology enables us to sidestep the same confounding factors while preserving the vignette's character, who is necessary to keep the semantic contradiction at the core of any scenario designed to compare DT and CHT.

> our interlocutors say in their utterances, enter into our inferences about what our interlocutors think and plan.
> (Cohnitz, 2015, p. 97)

> That meaning is consciously available to us, and thus in principle reportable, should arguably be taken as a fundamental fact that a theory of meaning has to accommodate.
> (Cohnitz and Haukioja, 2015, p. 634, n. 26)

Therefore, for example, Smith understands John's sentence "Federer is the best player in tennis history" through a subpersonal and implicit elaboration, but the result of that comprehension process is consciously graspable to Smith. He is aware of what John said and he now knows that John has a certain opinion about who the best tennis player in history is.

4.1.2 Are TVJs reliable comprehension reports?

If the previous arguments are correct, they show that the mere inclusion of a speaker in the prompt does not automatically entail overt theoretical reflection on the participant's part. However, one can preserve Martí's over-theorization criticism by framing it as stemming not from the presence of John or Ivy *per se*, but from their presence *in combination* with a TVJ task. One can argue that the TVJ formation and expression is a conscious process during which theoretical reflection may interpose, eventually influencing the resulting TVJ. In other words, one may agree that the comprehension process of a sentence that a third speaker utters is subpersonal and thus devoid of theoretical reflections; however, one matter is that comprehension process, while another matter is the TVJ formation subsequent to that. The participant may fail to form the TVJ on the (sole) basis of the proposition that she understood as expressed by John, and form it (also) on the basis of other theoretical factors. Notably, the epistemic ambiguity is one of those confounding factors. For instance, a participant may *understand* John's sentence "Gödel is the discoverer of the incompleteness theorem" as false, in line with CHT. Nonetheless, the same participant may *report* that the sentence is true because she forms that TVJ by reflecting on the speaker, and more specifically on the consensus within the

speaker's community and thus her blameless condition. Those epistemic considerations intervene during the process of the TVJ formation and taint the resulting TVJ.

My considerations follow the arguments that Cohnitz and Haukioja (2015, 2021) articulate. In the context of the experimental debate on proper names, scholars used the term "intuition" to label two different things: the interpretation of the name or of the name-containing sentence on the one hand, and the report of whom the name refers to or of the truth value of the proposition that the sentence expresses on the other. Cohnitz and Haukioja (2021) write:

> The metaphilosophical discussion of intuitions typically focuses on "intuitive judgements", that is, on a certain type of propositional attitude the content of which can linguistically be expressed by a that-clause. But this runs together the outputs of an intuitive capacity with the linguistic report of it. It is helpful to first distinguish intuitive capacities and the outputs of these capacities. We can do a lot of things intuitively: we can intuitively produce and interpret utterances, we can intuitively estimate the distance of an approaching vehicle, we can intuitively predict the trajectory of a thrown baseball.
> (Cohnitz and Haukioja, 2021, p. 556)[133]

Scholars have labeled those phenomena "intuitions" because they pertain to the spontaneous, that is intuitive, way in which a person interprets a word or a sentence, on the one hand, and in which she report her interpretation, on the other. One may question, as Devitt and Porot do (2018, p. 1579, n. 11), whether the word "intuition" is appropriate for both the facets that Cohnitz and Haukioja consider. Following Devitt and Porot themselves, during this dissertation I have used the word "intuition" only to label the judgment – as the one in MMNS (2004) – that a participant expresses as to whom John refers to when using the critical name. I have refrained from applying the term "intuition" when talking about the linguistic usage phenomena of production and comprehension. However, nothing conceptually relevant depends on the specific cases to which one chooses to apply the term "intuition", as Haukioja and Cohnitz themselves stress: «perhaps it would be better for the debate to eliminate that term. In any case, one can

[133] To be precise, Cohnitz and Haukioja (2015) distinguish four things that scholars happened to call "intuition". Although all four phenomena that Cohnitz and Haukioja target somehow intersect my arguments, I find it sufficient for my purposes to explicitly introduce and present only two of them.

understand our definitions as purely simulative ones, which use 'intuition' in lieu of a better word» (p. 622). Given those premises, what matters for our purposes is that a TVJ final prompt encompasses two different things: first, the interpretation of a sentence; second, the report – through a TVJ – of what the participant understood as expressed by that sentence. Theoretical reflection may intervene in the later phase, the one pertaining to the formulation of the TVJ concerning what John said. It is at *that stage* that Martí's concerns about the presence of a speaker may surface. Some participants may express their TVJ because they start reasoning on how John uses the name "Gödel". They may reflect on what makes it possible for John's use of the name to pick out an extra-linguistic entity. Therefore, participants may ultimately choose the explanation – whether descriptivist or causal-historical – that they find theoretically more intuitive. Finally, they may select the TVJ based on that explanation rather than on the proposition that they understood as expressed by John's claim. Thus, the TVJ fails to reveal the proposition that the participant understood, and rather sheds light on how the participant thinks that people use proper names.

One may wonder how plausible it is that such a confounding factor, which revolves around a reflection on the name "Gödel" as John uses it, intervenes during the participant's TVJ formation process. Arguably, that theoretical confound is less likely to emerge in the context of a TVJ prompt when compared to RI one. As said, a TVJ final prompt, in contrast to RI one, does not explicitly ask anything about how John uses the name "Gödel" or whom John talks about when uttering it. Domaneschi and Vignolo (2020) stress that point when they comment on the implications of the TVJ "True" – epistemic ambiguity aside – in response to a sentence like "Tsu Ch'ung Chih was a great astronomer":

> Indeed, there is no evidence that the participants [with the TVJ "True"] make the metalinguistic judgment that Ivy refers to X [the real discoverer] when she uses the name "Tsu Ch'ung Chih". What their response "true" evidences is that they understand Ivy as referring to X, not that they judge that she refers to X with the

> name "Tsu Ch'ung Chih." Likewise, if participants choose "false," then this is evidence that they understand Ivy as referring to Tsu Ch'ung Chih [the thief]. (Vignolo and Domaneschi, 2020, p. 445)[134]

Furthermore, the theoretical confound under discussion is arguably less likely to emerge than the epistemic-ambiguity one. The latter regards a concept – the truth concept – that the participant directly and consciously employs while formulating the TVJ. Instead, participants run into the kind of theoretical reflection that Martí warns against if they engage in a theoretical examination of how a certain speaker, like Ivy, uses the name "Gödel". However, as said, a TVJ final prompt does not explicitly ask anything about how a speaker uses the name "Gödel" or whom she talks about: it asks whether a certain sentence is true or false. In light of that, it is arguably less plausible that a participant starts a theoretical reflection concerning how John uses the name rather than one concerning the perspective from which to interpret the truth predicate.

Despite the foregoing concerns, one can insist that some participants may possibly perform a TVJ final prompt in the way that Martí describes, namely, by grounding the TVJ on a theoretical reflection on the mode of reference of "Gödel" as John uses it. After all, the Gödel Case presents a pretty intricate scenario, and that complexity may incline participants toward a deliberate analysis of the single words of the sentence that they have to judge. Moreover, the experimental setting itself may prompt the participant to embark on reflections she would not typically engage in during a natural and actual conversation – such as one with a friend. Therefore, while theoretical reflections on the names' mode of reference may contaminate a TVJ arguably in a more conjectural way when contrasted with a RI, it is nonetheless important to explore alternative ways to test linguistic comprehension.

4.1.3 Eye-tracking and the sensitivity to the others' mental states

Cohnitz and Haukioja (2015) suggest that a way to overcome the potential theoretical nature of TVJs comes from psycholinguistics, where scholars use the eye-tracking methodology to test linguistic comprehension. Within an eye-tracking study employing a

[134] Analogous considerations emerge in Vignolo and Domaneschi (2018, p. 4). Note however that Domaneschi (personal communication) has developed a growing sense of caution on the issue and thinks that the participant's potential reflections on how John uses "Gödel" that Martí emphasizes might be a real cause of concern also for the evaluation of TVJs.

visual word paradigm, the subject is presented with images depicting the potential referents of the words that she is hearing. Simultaneously, the eye movements are monitored. In this way, researchers can investigate the referential connections that the participant establishes between the words she listens to and the referents she sees pictured in front of her (Tanenhaus and Spivey-Knowlton, 1996). The crucial advantage of this methodology lies in its capacity to obviate the need for participants to express any explicit judgment, thereby eliminating the potential influence of theoretical reflection. Moreover, conscious theoretical reflections cannot contaminate eye movements, as the latter are natural, spontaneous and involuntary. In other words, TVJ is an offline methodology since the measure that is supposed to unveil the participant's linguistic comprehension appears *after* the comprehension process itself. The subject *first* reads John's sentence and *then* evaluates it for its truth value: during the gap between the two stages, theoretical reflections may intervene, finally influencing the TVJ itself. Instead, eye-tracking is an online methodology because it sheds light on the participant's linguistic comprehension *while* the understanding process unfolds, that is, while the participant listens to John's critical sentence. Since – as argued – comprehension is largely a subpersonal and implicit task, the eye-tracking methodology minimizes the risk that explicit theoretical reflections may surface and compromise the resulting data. To use Tanenhaus and Spivey-Knowlton's (1996) words: «spoken language processing [thanks to eye-tracking] can be studied both with real-time precision and with a non-invasive task – that is, a task that does not require interrupting the speech stream or *having the subject make an overt decision about the speech itself*» (p. 586, my emphasis). Furthermore, the eye-tracking methodology allows us to bypass the epistemic ambiguity inherent in TVJ tests, since it does not require participants to employ the truth predicate at all.

As examples of eye-tracking works exploring how participants disambiguate reference, Cohnitz (2015) mentions Karabanov et al. (2007) and Keysar, Lin and Barr (2003). To complement this list, I include the studies of Altmann and Kamide (1999) and Ferguson and Breheny (2011). In this section, I present Ferguson and Breheny's experiment in some detail because it sheds light upon a methodological aspect that holds relevance for any potential eye-tracking study that aims to test the Gödel-Schmidt scenario. Ferguson and Breheny's work locates itself in the context of the literature on the so-called Theory of Mind, namely, «the ability to understand and predict events in

terms of other people's mental states, such as their intentions, beliefs and desires» (p. 179). They test the hypothesis that the participants' spontaneous comprehension process makes use of the mental contents that participants ascribe to others, both when those are first-level mental states and when they are higher-order ones. For example, John's basic preference toward a specific color or food is a first-level mental content because it involves only John and his preference toward a specific food or color. Instead, John's desire that no one knows that a specific color or food is his own favorite is a higher-order mental content because it is a desire that Tom develops based on the mental state that he would like other people to hold about his own food or color preferences.

In their Experiment 1, Ferguson and Breheny assign participants to one of two conditions: the *open condition* and the *secret condition*. In both, participants first listen to a sentence like: "Tom's favorite color is pink". The two conditions differ as to how the auditory stimulus proceeds. In the open condition, participants hear "Tom is always telling people that his favorite color is pink. Last week Tom bought a new car and he deliberately chose a pink car". Instead, in the secret condition, participants hear "Tom does not want anyone to know that his favorite color is pink. Last week Tom bought a new car and he deliberately chose a green car". Table 4.1 reports the entire auditory stimuli of the two conditions.

Table 4.1 Auditory stimuli in Ferguson and Breheny's Experiment 1

Condition	**Auditory stimulus**
Open	Tom's favorite color is pink. Tom is always telling people that his favorite color is pink. Last week Tom bought a new car and he deliberately chose a pink car.
Close	Tom's favorite color is pink. Tom does not want anyone to know that his favorite color is pink. Last week Tom bought a new car and he deliberately chose a green car.

While the participant listens to the last sentence (the one starting with "Last week Tom"), she is presented with a screen portraying four visual stimuli, and her eye movements are tracked. The visual stimuli consist of a pink car, a green one and two

distractors. As expected, when the participants in the open condition listen to "pink car" and those in the secret condition listen to "green car", they clearly direct their gaze toward the pink car and the green car, respectively. However, interestingly, the differentiation in eye movements between the two groups emerges even before the disambiguating phrase "pink/green car". Participants in the open condition and those in the secret start exhibiting a partial and initial tendency to look at the pink and green car respectively when they hear "car" in "Last week Tom bought a new car".[135] Participants can anticipate the sentence referent before hearing the disambiguating phrase because their real-time sentence-comprehension process is sensitive to Tom's mental contents. Specifically, participants process the sentence according to Tom's higher-order desire, because one naturally expects Tom to ground his purchase choice on that. If Tom does not want anyone to know that his favorite color is pink, he will not buy a pink car because that choice would reveal his color preference (secret condition). If Tom always tells anyone that his favorite color is pink, he has no problem letting other people know of that preference and therefore he will buy a pink car (open condition). These inferences, grounded in Tom's mental contents, shape the participants' real-time understanding processes and eye movements. Therefore, the findings of Experiment 1 suggest that people's real-time sentence comprehension, albeit devoid of conscious reflection, involves inferences rooted in the mental contents of others.

[135] Some subtle divergences emerge across the two conditions as to the exact moment in which participants start exhibiting a tendency to fixate their gaze upon one specific car, and as to how that tendency decreases or increases during the entire auditory stimulus. However, those differences do not impact Ferguson and Breheny's main finding, namely, that participants anticipate the sentence referent.

Ferguson and Breheny conduct also another study, Experiment 2. Compared to Experiment 1, the main difference is the final sentence of the auditory stimuli across the two conditions. Table 4.2 reports an example.

Table 4.2 Auditory stimuli in Ferguson and Breheny's Experiment 2

Condition	**Auditory stimulus**
Open	Tom's favorite color is pink. Tom is always telling people that his favorite color is pink. Last week Tom bought a new car and he surprisingly chose a green car.
Close	Tom's favorite color is pink. Tom does not want anyone to know that his favorite color is pink. Last week Tom bought a new car and he surprisingly chose a pink car.

The auditory stimulus introduces an event that is indeed "surprising", as it contradicts the higher-order desire that one would naturally expect to guide Tom's choice of the car. Ferguson and Breheny's rationale is to outline a scenario in which Tom sets his higher-order intention aside and chooses to ground his purchase upon his basic preference. Therefore, the aim of Experiment 2 is to ascertain whether the participant's sensitivity to Tom's mental contents can easily shift from the higher-order ones to the basic ones upon encountering an unexpected and indeed surprising choice.

Also in this case, Ferguson and Breheny find that participants can anticipate the referent before listening to the disambiguating phrase: even before hearing "pink car" or "green car", they start displaying an initial and partial tendency to look at the appropriate referent – the pink car in the secret condition and the green one in the open condition. More specifically, that inclination emerges at the earliest possibility, namely, as soon as the sentence hints at an impending surprising event, i.e. from the onset of the adverb "surprisingly". That result underscores the participants' ability, during their online comprehension process, to seamlessly integrate the mental states that they ascribe to the character, in this case Tom, with the «local lexical information», in this case the adverb "surprisingly" (p. 193). Participants are aware of Tom's basic preference and his higher-order intention, and one would expect Tom to ground her choice on the latter. However,

as soon as participants listen to the adverb "surprisingly", their linguistic comprehension of the sentence proves to be sensitive to Tom's basic preference and their eye movements anticipate the sentence referent based on that.

For our purposes, the finding that participants are able to anticipate the referent, while very interesting, is not crucial *per se*. What is central for our goals is that the reference anticipation reveals that the participants' eye movements during the sentence-comprehension process are sensitive to the mental contents that the vignette ascribes to the character – that is, Tom. In other words, participants vividly track and respond to the mental contents that the vignette ascribes to a specific character and those contents impact the participant's online language comprehension, as evident from their eye movements: «ToM inferences can be made online and very rapidly indeed» (p. 188). Interestingly, participants are so sensitive to such contents as to keep track of both the character's basic preferences and of her higher-order ones: «both conditions involve fairly sophisticated ToM reasoning, as characters make choices based not only on their basic preferences but also on how they wish to be seen in the eyes of others» (p. 182). Such a task may seem cognitively demanding. However, the experiments reveal that the participant's real-time, and thus not consciously reflective, sentence processing is remarkably attuned to the mental contents that the vignette ascribes to Tom. The eye movements anticipate the contextually appropriate referent *precisely because* the participant's comprehension process keeps track of those mental states. To use Ferguson and Breheny's words, the two studies «have identified contexts where even complex higher-order ToM inferences can be generated rapidly and online, as the concurrent discourse is unfolding» (p. 194). In other words, «language comprehension is a truly interactive process that immediately makes use of available information from both prior context and local semantics» (p. 192).

At this point, it is possible to identify the correlation between Ferguson and Breheny's results and our eye-tracking study on the Gödel Case, which I am about to present in the forthcoming section. I have argued that Ferguson and Breheny's study corroborates the thesis that the subject's eye movements are sensitive to the mental contents that the vignette attributes to the relevant characters of the scenario. If that interpretation is correct, then, in our experiment, the eye movements of the descriptivist – during her comprehension of the name "Gödel" – should be sensitive to the belief that John has in relation to the name "Gödel", namely, that he is the discoverer of the incompleteness

theorem. In other words, also in our study the participant's eye movements should be responsive to both the «local semantics», namely the fact that John is using a proper name, and to the «prior context», namely the belief that John associates with the name (p. 192). As pointed out several times, the sensitivity toward the information that John, namely the speaker in semantic contradiction, associates with the name is crucial to compare DT and CHT. The divergence between the two theories can manifest within an eye-tracking design only if the participant keeps track of John's belief and thus her eye movements – if she is a descriptivist – reflect a comprehension of the name based on that belief. Ferguson and Breheny's study seems to provide solid grounds to argue that the above methodological assumption is valid within an eye-tracking study testing the Gödel Case.

4.2 Method, material, participants and hypotheses

Filippo Domaneschi, Massimiliano Vignolo, Camilo Rodríguez Ronderos and I coauthored an eye-tracking experimental study on the Gödel Case.[136] We tested 50 Italian native speakers (M = 25.50; SD = 3.52; 31 females; 19 males) by using 10 Gödel-style vignettes (randomized and in Italian). Figure 4.1 reports one item as a way of example. Participants read the first part of the vignette (*Step 1*), which introduces the causal-historical referent (which we call "Target") and a picture of it (*Step 2*). Participants then read the second part of the vignette (*Step 3*), which introduces the descriptivist referent ("Competitor") and a picture of it (*Step 4*). Participants read the final part of the vignette (*Step 5*). After a 2-second interruption (*Step 6*), four pictures (Target, Competitor and two distractors) appear on the screen (*Step* 7), randomly distributed in the four angles of the screen. Eye movements are recorded using a Tobii Pro X3-120 eye-tracker, while participants listen to the sentence that a speaker in semantic contradiction utters. The voice uttering the critical sentence is a male one if the speaker is a man while a female one if the speaker is a woman: that choice intends to further prompt the participant's comprehension process to elaborate that sentence *as* one that the speaker in semantic contradiction utters. The speaker's claim has either a name-description order or a

[136] We tested (only) Gödel-style vignettes, instead of using (also) Jonah-style one(s), for two reasons. First, most of the literature focuses on the Gödel Case, making it more conducive to meaningful comparisons. Second, at least *prima facie*, applying the eye-tracking technique to test a Jonah scenario is challenging. The methodology requires two different competing visual referents, yet the crucial point of a Jonah Case is precisely that, under one interpretation of the name (the descriptivist one), the name lacks a referent. Hence, it is not possible to present the participant with a visual stimulus corresponding to that reading of the name. To use Tanenhaus and Spivey-Knowlton's (1996) words, with eye-tracking «only language that makes reference to concrete objects can be used» (p. 586).

description-name order, in a within-subjects design: across the set of the ten items, each participant receives a random selection of items in one order and the remaining ones in the alternate order. In the former condition, called "FIRST", the sentence is as follows: "Carlo Model is the discoverer of the Infinitesimal Theorem". In the latter, called "LAST", the sentence is as follows: "The discoverer of the Infinitesimal Theorem is Carlo Model". The rationale is to verify whether, while hearing the proper name, participants look at Target or at Competitor, in different syntactic positions. After listening to the sentence, the participant is asked to express her truth-value judgment by answering this question: "…do you believe that his/her [the vignette's character] claim is true or false?" (*Step 8*). All the materials are available on OSF.[137] We introduce the TVJ final prompt because our study not only aims to test comprehension by collecting eye-tracking data, but also intends to investigate the dependability of TVJs. For the reasons that I have presented, eye-tracking constitutes the benchmark technique to test linguistic comprehension, namely, the methodology that provides the most reliable data. Therefore, an alternative methodology, like TVJs, aimed to test the same phenomenon, is reliable to the extent that it yields the same results as the ones deriving from the eye-tracking technique.[138]

[137] At this link: <https://osf.io/kgnte/?view_only=4bcf12d5c2444620a89ec9d2129da945>. Unfortunately, our vignettes' wording inherits a bias against DT from the previous literature. For example, the vignette states that "Like any other student, Maurizio believes that Carlo Model is the author of an important mathematical theorem, called 'Infinitesimal Theorem'". For the reasons that I have discussed (section 2.2.2.3 and 2.2.2.4), a descriptivist narrator could not utter that claim, unless she is taking a *de dicto* mode of presentation of Maurizio's mental contents.

[138] That rationale follows a suggestion deriving from Cohnitz and Haukioja:

> Since these eye movements are automatic, not under our voluntary control, and clearly correlated with language comprehension, this data provides evidence about how hearers semantically interpret referential expressions that is much more direct than that obtainable from surveys. There is no need to ask the test subjects afterwards how they interpreted the utterance semantically, *but if one would like to know how reliable we are in reporting such first-level comprehension intuitions [i.e., TVJs], one should test whether these reports agree with the eye-tracking data*.
> (Cohnitz and Haukioja, 2015, p. 639, my emphasis)

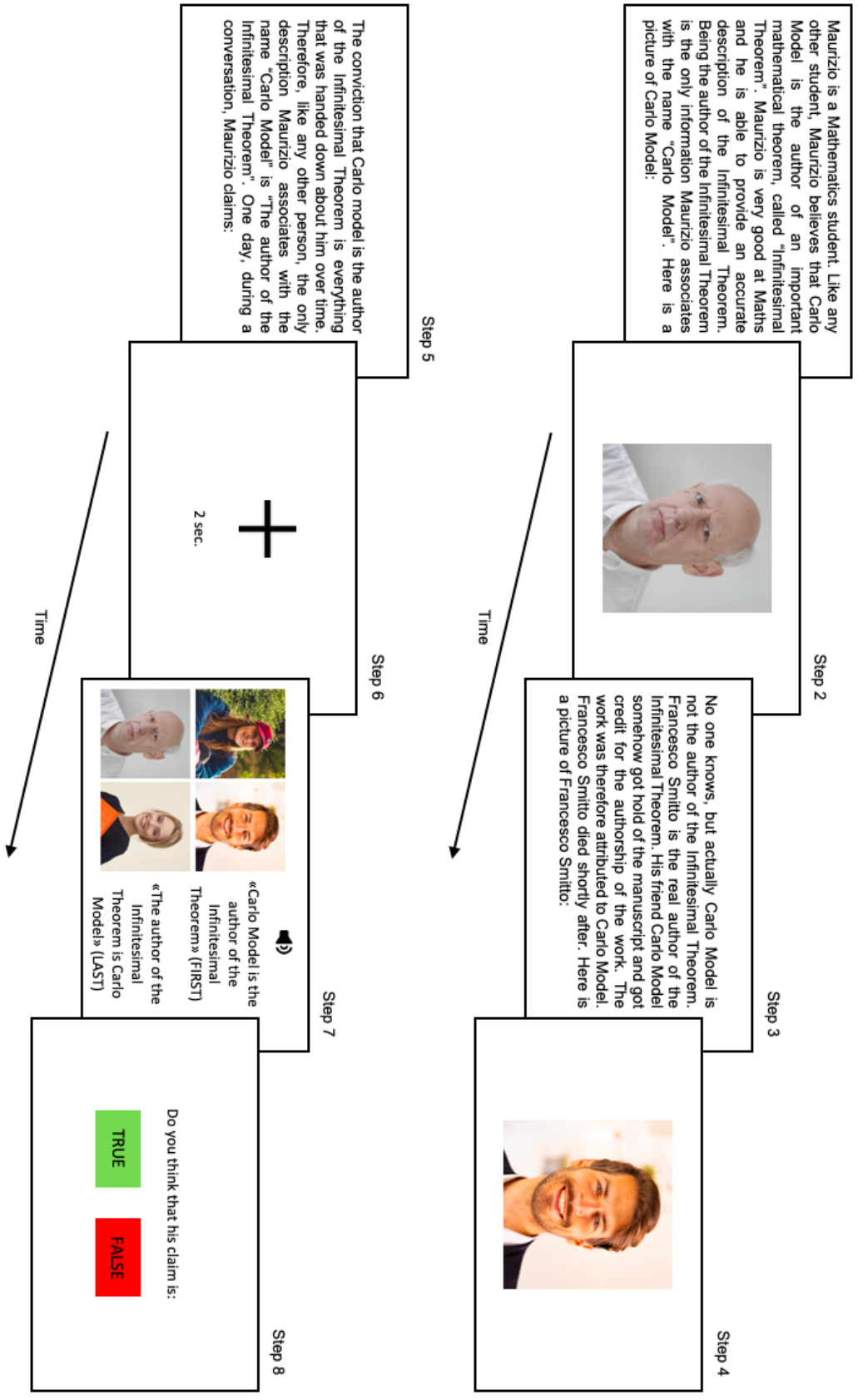

Figure 4.1 Example of vignette (translated into English) of Experiment 3

Our predictions are straightforward. If a participant is a causal-historical user of proper names, then, when listening to the name "Carlo Model" that Maurizio utters, she will look at Target, namely, at the referent that CHT predicts. If instead the participant is a descriptivist, then, when listening to that name, she will look at Competitor, namely at the referent that DT predicts. In other words, the participant will look at the individual whom she understands as the name's referent.

Prior to delving into the presentation and subsequent analysis of our findings, it is important to stress an important difference between the way Cohnitz and Haukioja (2015) frame the advantageous impact of the eye-tracking technique and the way we implement it in our experiment. They write:

> We believe that this [the eye-tracking methodology] is more likely to be a way forward in the experimental investigation of semantics than the use of vignettes modelled after thought experiments […]. Perhaps there are other experimental methods that could be used, too. *The main shortcoming of the existing empirical studies, we think, is that they have not distanced themselves far enough from the thought experiment setups commonly used in philosophy.* As we have explained above, thought experimentation about fictional (and often far-fetched) scenarios can be a perfectly legitimate tool for a philosopher to investigate his or her own […] intuitions, but when we study non-philosophers, who lack expertise in dealing with thought experiments, other methods may be in order.
> (Cohnitz and Haukioja, 2015, p. 639, my emphasis)

Cohnitz and Haukioja argue that one of the problems that the literature has faced is the use of vignettes modeled upon philosophical far-fetched thought experiments, like the Gödel Case, and the eye-tracking methodology would serve as a means to circumvent that dependence. We do not agree with that, and indeed our eye-tracking study employs versions of the classical Gödel scenario. Regardless of the chosen measurement technique (RIs, TVJs, eye-tracking and so on), the scenario that cases like the Gödel one describe remains indispensable to elicit data suitable to compare DT and CHT. It is essential to delineate a story in which a community of speakers associates the name with a description that picks out an individual being different from the one to whom the name's causal-

historical chain leads – that is precisely the point of disagreement between the two theories. That said, Cohnitz and Haukioja are right that the Gödel scenario exhibits a certain degree of complexity. We also believe, though, that the eye-tracking methodology can indirectly help participants to partially cope with the vignette's intricacies. An eye-tracking design monitors only the participant's involuntary and spontaneous eye movements, and thus relieves her of the duty of formulating a judgment (whether it be a RI or a TVJ). This approach reduces the overall cognitive burden that the subject has to manage while addressing a vignette. The participant can allocate her entire cognitive faculty to grasping the scenario and the critical sentence, without the added task of formulating a judgment regarding that sentence.[139]

4.3 Analysis

To analyze the truth-value judgments, we calculated the percentages of "True" and "False" responses across participants in the two conditions (FIRST and LAST). We then fitted a mixed effects, logistic regression model to the data, including random intercepts and slopes by items and by participants. To analyze the eye-tracking record, we calculated the proportions of looks to Target, Competitor, and distractors while participants listened to the sentences. We included only those items for which participants answered "True" in the TVJ task. Proportions of looks were calculated for every 20 milliseconds time bin across two pre-determined time-windows: (i) the "critical" region, time-locked from the onset until the offset of the critical word (the proper name), and (ii) the "description" region, time-locked from onset to offset of the description (i.e., the remaining information in the sentence). We then fitted mixed effects linear regression models to each time window, including random intercepts and slopes by items and by participants. The dependent measure used was log-gaze probabilities of looks to Target divided by looks to Competitor (Arai et al., 2007). This measure allows us to quantify whether there is a significant preference to look at either Target or Competitor in each of the time-windows. The models had the factor 'Position' as a predictor, which was treatment contrast coded. We fitted each model twice, changing the condition coded as the baseline (either FIRST or LAST). This was critical to test whether the intercept of the model was significantly

[139] As said, each vignette of our eye-tracking experiment includes a TVJ task, which however is subsequent to the collection of the eye-tracking data for that vignette. Therefore, the cognitive effort necessary to the TVJ task does not impact on the previous eye-tracking phase for that vignette.

different form zero in each case, which amounts to testing whether there was a preference for looking at Target (positive numbers) or Competitor (negative numbers) in each of the conditions.

4.4 Results

The regression model for the truth-value judgment task failed to find a statistically significant difference between the truth-value judgments across the two conditions (z = 1.29, $p = 0.19$) (Figure 4.2). In both conditions, about 85% of participants answer that the sentence is true: 86% in FIRST and 84% in LAST.

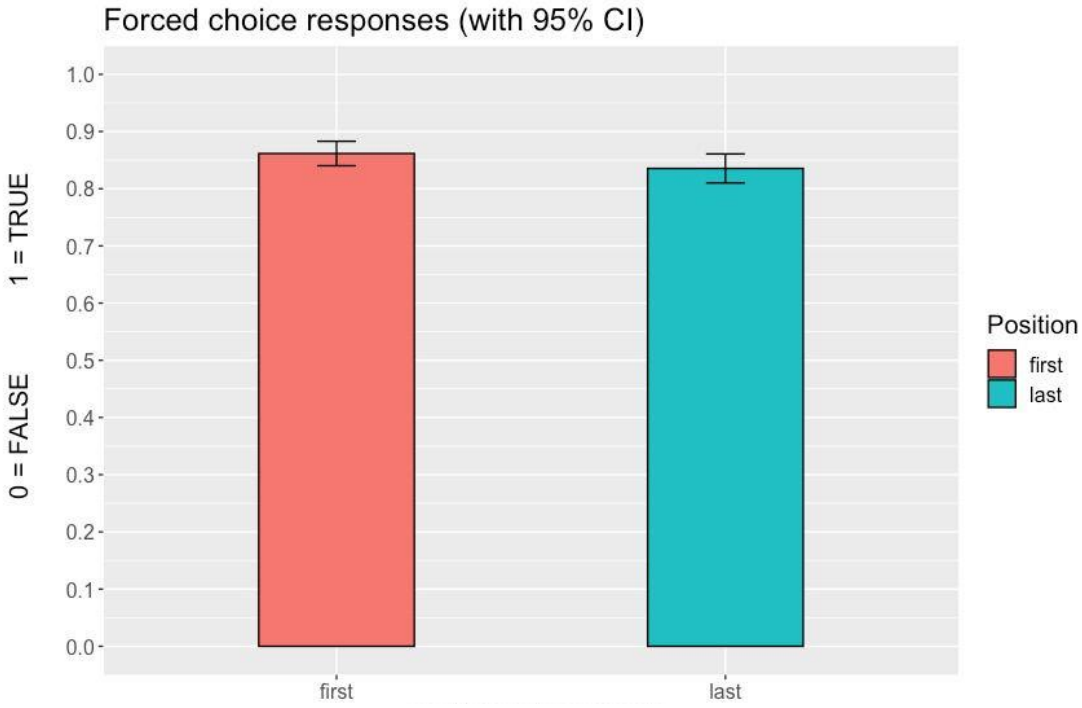

Figure 4.2 Proportions of "True" / "False" answers to the TVJ task in Experiment 3

Figure 4.3 and Figure 4.4 show the results of the eye-tracking analysis, calculated only for those items for which participants answered "True" in the TVJ task. The model for the critical region (Figure 4.3) showed a significant difference between conditions (t = 3.573, ß = -0.46133, $p < 0.001$). Re-fitting the model with different baseline conditions showed that FIRST condition (where the name precedes the description) was positive and significantly different from zero (t = 6, ß= 0.35, $p < 0.001$), while LAST condition (where

the description precedes the name) was negative and not significantly different from zero (t = 0.9, ß = -0.1, *p* =. 0.3).

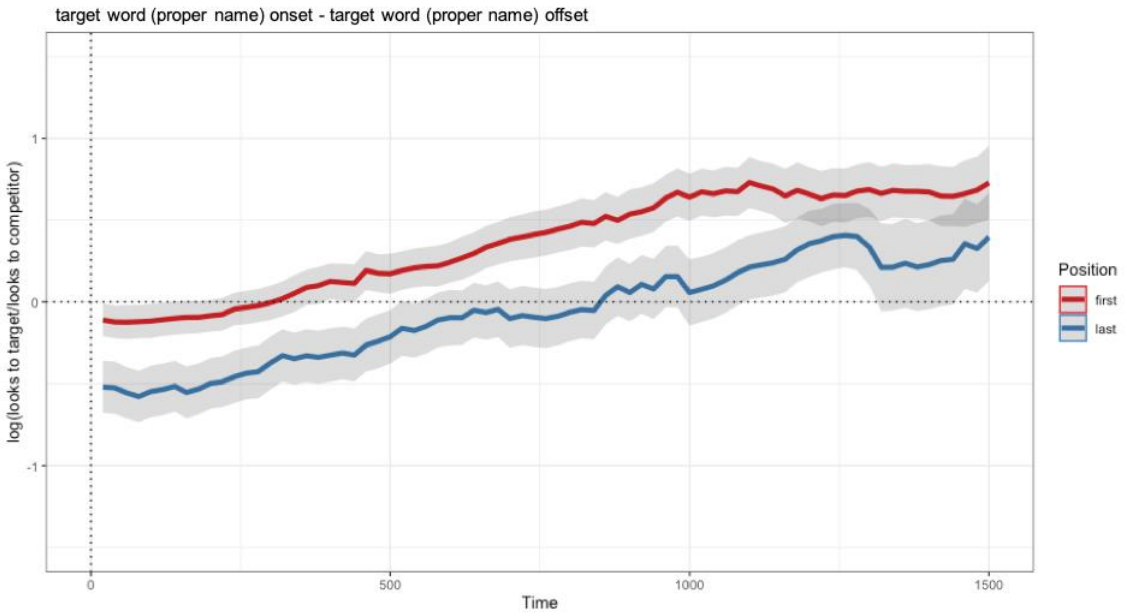

Figure 4.3 Log-gaze probabilities of looks to Target divided by looks to Competitor, time-locked from onset to offset of the proper name time-window in Experiment 3

The model for the 'description' region (Figure 4.4) also showed a significant difference between conditions (t = 7.3, ß = 0.98, $p < 0.001$). Re-fitting the model with different baseline conditions showed that FIRST condition was positive and significantly

different from zero (t = 7.8, ß= 0.6, *p.* < 0.001), while LAST condition was negative and significantly different from zero (t = 3.8, ß. = -0.3, *p* < 0.001).

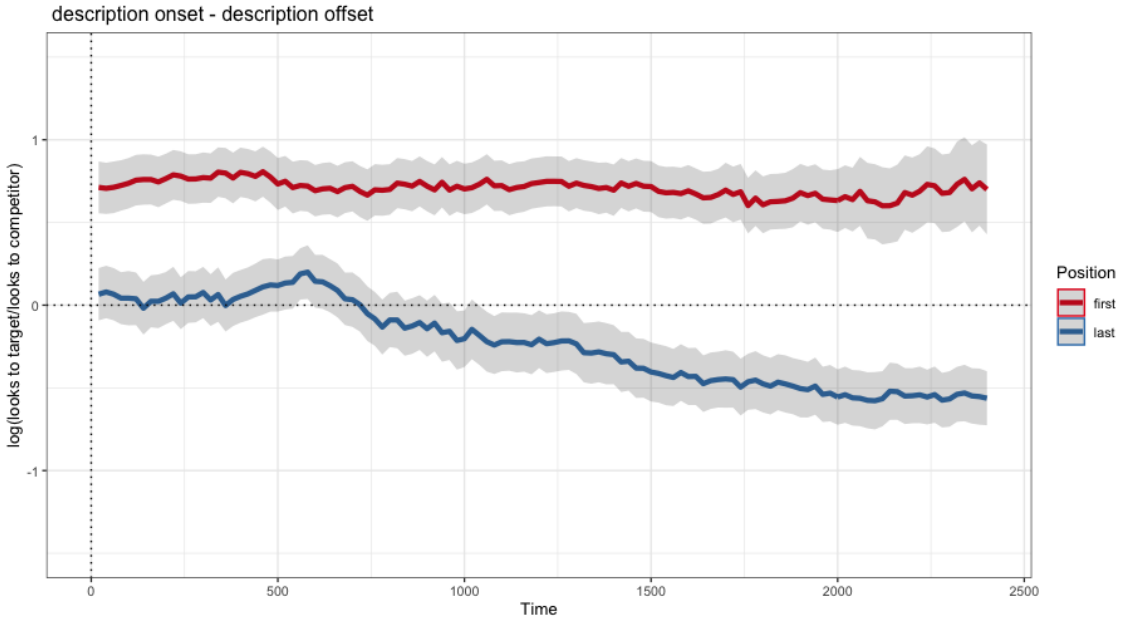

Figure 4.4 Log-gaze probabilities of looks to Target divided by looks to Competitor, time-locked from onset to offset of the description time-window in Experiment 3

4.5 Discussion

The previous literature on the Gödel Case with Western and Italian participants signals a degree of variation as to TVJs: for example, Li's (2021) American sample in her Original Condition chooses the Kripkean TVJ "False" 90% of the time, while Li et al.'s (2018) American sample 65% of the time – thus constituting a noteworthy disparity of 25 percentage points. However, even considering that variation, our TVJ results remarkably deviate from the existing literature in two respects. First, the previous experiments showed a consistent proclivity to answer that the sentence is false, as CHT predicts: for example, if one focuses on the two extant studies investigating TVJs with Italian participants, namely Vignolo and Domaneschi (2018) and Domaneschi and Vignolo (2020), 65% and 58% of participants opt for the TVJ "False", respectively. Our sample, instead, chooses the TVJ "True" approximately 85% of the time, in both conditions.[140]

[140] Moreover, it is noteworthy that, in the existing literature, not even Chinese samples, who proved to be inclined to choose the purportedly descriptivist TVJ, ever reached a proportion of TVJs "True" as high as 85%. Li's (2021) Experiment 1 provides the Chinese highest proportion of TVJs "True", which peaked at 68%.

Second, as discussed, some confounding factors present themselves as plausible explanations of the variation within the previous literature: the attributive understanding of the name and the social interpretation of the predicate. However, our vignettes control for those factors and, when that happens, the proportion of Kripkean TVJs typically increases (Experiment 1 in Li, 2021; Study 3 in Li, 2023a; my Experiment 2). In light of this, we are left without a clear explanation for why our scenarios elicit the vast majority of participants to answer that the sentence is true, as DT predicts.[141] Our design introduces some peculiar variables that may have somehow influenced the results, such as the use of pictures[142] or the fact that we present the critical sentence via an auditory stimulus rather than a written one. While these factors may have impacted the results, the precise mechanisms at play remain blurry. In any case, the collected TVJs seem to reinforce the case of those philosophers who take the experimental data on the Gödel Case to constitute a threat to Kripke's refutation of descriptivism. Therefore, it becomes even methodologically more paramount to analyze the eye-tracking data, which provides the gold standard in discerning how participants indeed understand the critical sentence and the proper name therein.

As mentioned, we calculated the eye-tracking data only for those items for which participants provided the TVJ "True" (as just seen, participants opt for that TVJ a remarkable majority of the time). This selective approach stems from the consideration that the answer "True" is supposed to provide evidence against Kripke's refutation of classical descriptivism and has also been taken to be affected by the ambiguity of the truth predicate. The eye-tracking data serves as a valuable instrument for gauging the extent to which those participants express a genuine descriptivist understanding of the name.

In FIRST condition, participants listen – for instance – to Maurizio's claim "Carlo Model is the author of the Infinitesimal Theorem". As to the first part of the sentence, that is the proper name "Carlo Model", the result is straightforward: participants fixate their gaze upon Target, as evident from the red line in Figure 4.3, which is positioned above the neutral line in a statistically significant way. Interestingly, the participants

[141] Especially in the light of the eye-tracking data, which – as I am about to show – support CHT as the correct theory of reference.

[142] The use of pictures it is not an absolute peculiarity of our study, as in Li et al. (2018) too «the stories are accompanied by clipart pictures to engage participants» (p. 109). The incorporation of pictures in Li et al.'s study is due to the inclusion of also children as participants, who may encounter more difficulties than adults in following the Gödel-style stories if presented solely in a textual format.

continue looking at Target even when they progress to the second segment of the sentence, namely the definite description "The author of the Infinitesimal Theorem", as evident from the red line in Figure 4.4, which continues being positioned above the neutral line in a statistically significant way. The outcome regarding the second part of the sentence may appear counterintuitive, as the satisfier of the description is Competitor, namely Francesco Smitto, and thus one may expect the participants, when hearing the description, to redirect their gaze upon him. A plausible explanation for the continuous fixation on Target relies on the fact that the proper name takes the grammatical subject position within the sentence structure. As a consequence, participants keep their gaze on the individual corresponding to the subject, namely Target, because that is the individual to whom the sentence ascribes a property, namely that of being the author of the theorem. In other words, participants understand the speaker to use the definite description for predicating of Target the property of being the author of the Infinitesimal Theorem.

In LAST condition, participants listen to the sentence "The author of the Infinitesimal Theorem is Carlo Model". As to the first part of the sentence, namely the definite description "The author of the Infinitesimal Theorem", participants fixate their gaze upon Competitor, as evident from the blue line in Figure 4.4, which is positioned below the neutral line in a statistically significant way. This is an expected finding because Competitor, namely Carlo Smitto, is the satisfier of the description. However, the situation becomes more intricate when one examines the data regarding the second segment of the sentence, namely the proper name. Participants start exhibiting a different trend: their gaze shifts from Competitor to Target, as evident from the blue line in Figure 4.3, which transitions from below the neutral line to above it. However, the model represented by the blue line remains overall negative (that is, generally below the neutral line), although in a statistically non-significant way.[143] Given that Target is the referent that CHT predicts, one may argue the above circumstance is a cause of concern for the Kripkean theorist. We do not think it ultimately is, though, for two reasons.

First, although the overall trend is negative, that is not statistically significantly so; moreover, a noticeable pattern emerges: upon hearing the proper name, the participants' gaze stops being statistically significantly directed toward Competitor and starts moving

[143] Indeed, the eye-tracking results as regards the proper name in FIRST condition differ statically significantly from those as regards the name in LAST.

toward Target, namely, the individual that CHT predicts. It seems reasonable to conjecture that the overall trend does not become positive and statistically significant because the auditory length of the name does not afford participants enough time to completely redirect their gaze toward Target.

Second, like in FIRST condition, also participants in LAST may retain some lingering inclination to continue looking at the individual who corresponds to the grammatical subject of the sentence. Since in LAST the grammatical subject is the definite description, participants partially persist in looking at Competitor because that is the satisfier of the description and thus the individual to whom the sentence ascribes a property, namely, that of being identical to Carlo Model.

4.6 Conclusions

Our eye-tracking data provide evidence against DT. When participants hear the proper name, they tend to look at Target, namely the individual that CHT predicts as the name's referent. We have also argued that, even when that trend is not positive and statistically significant, that circumstance should not trouble the Kripkean theorist. With regard to the reliability of TVJs, reaching firm conclusions is more challenging. If one considers exclusively the TVJs of our eye-tracking experiment, they appear to be highly unreliable. If TVJs were a dependable way to test theories of reference and if we are right that eye-tracking data constitute the primary benchmark to study linguistic comprehension, then participants' TVJs should align with the eye-tracking results, and thus the Kripkean prediction. That is not what we have found, though, as only a small minority of participants – i.e., 15%, in both conditions – choose the TVJ "False", as per CHT. However, my Experiment 2 corroborates the previous tendency (e.g., Experiment 1 in Li, 2021) that Gödel-style vignettes with Western participants, when controlling for the attributive use of names and the potential social interpretation of predicates, elicit a very high proportion of causal-historical TVJs, with only a small minority of purportedly descriptivist TVJs. As discussed, that minority likely stems from the adoption of the community's perspective. Therefore, Experiment 2 and the previous studies suggest that Westerners provide overall dependable TVJs, that is, TVJs that properly shed light on the comprehension side of language: the limited extent to which those tests are unreliable coincides with the extent to which participants adopt the community's perspective. As

suggested, the exceptional TVJ results in the context of our eye-tracking experiment may be due to some specific features of our setting. For example, one may speculate that the circumstance that participants listen to John's voice uttering the critical sentence may somehow strengthen the participant's potential inclination to adopt the community's perspective and thus express the purportedly descriptivist TVJ. Alternatively, following Martí's theses, one may suggest that TVJ prompts face not only an epistemic ambiguity but also the possibility that participants meta-linguistically reflect on how John uses proper names. That is, our design would somehow strengthen the participants' potential inclination to such reflections. In any case, the considerably high proportion of purportedly descriptivist TVJs in our Experiment 3 suggests that the dependability on TVJs can oscillate due to some design-related factors that still lack precise identification. Hence, the findings from Experiment 3 caution against overreliance on TVJs as a means to test theories of reference, thereby at least partially substantiating Martí's reservations about the TVJ methodology.[144]

[144] At least in the context of Gödel-style and Jonah-style cases. In other contexts that delineate less far-fetched scenarios and do not present any epistemic ambiguity, the experimental research may likely place higher confidence in TVJs.

5 Conclusions

Chapter 1 constitutes the introductory part of my thesis. I present the two main views on the reference of proper names: the descriptivist theory and the causal-historical theory. In *Naming and Necessity*, Kripke grounds his arguments on some thought experiments, the most famous of which is the so-called Gödel Case. Machery, Mallon, Nichols and Stich, in 2004, publish the ground-breaking article "Semantics, cross-cultural style", in which they test the Gödel Case and obtain results that unveil a significant cross and intra-cultural variation in RIs. Therefore, Kripke's supposed refutation of descriptivism comes under attack. MMNS's study face a number of criticisms, though. Some of those accept that RIs are suitable data to test theories of reference, and yet emphasize that MMNS's final prompt is somehow ambiguous (Deutsch, 2009; Ludwig, 2009; Sytsma and Livengood, 2011). Some other criticisms, instead, question that RIs are the kind of data against which scholars should test theories of reference. Deutsch (2009, 2010, 2015) and Martí (2009, 2012, 2014, 2020) claim that RIs are not the evidence upon which Kripke himself relies while developing his Gödel Case. In line with the reasons that Devitt (2015c, 2020) provides, I do not share Deutsch and Martí's interpretative thesis. However, beyond that hermeneutical issue, the common core of this second line of objection – a core that Devitt himself endorses (2011, 2012b, 2015b) – is that scholars should test theories of reference against linguistic usage and not intuitions. Linguistic usage comprises production and comprehension. A way to investigate the latter consists in TVJs, which however suffer from two confounding factors: an epistemic ambiguity (Domaneschi and Vignolo, 2020; Li, 2021) and a possible attributive reading of names (Heck, 2018; Vignolo and Domaneschi, 2022). Finally, I present the comparatively less extensive literature on the Jonah Case, showing where its limitations lie and the necessity for further investigation.

In Chapter 2, I defend the thesis that scholars need to test theories of reference against linguistic usage, and thus I analyze Devitt and Porot (2018), as a paradigmatic example of a linguistic-production study on proper names. Devitt and Porot's criticisms against the formulation of MMNS's vignettes are unfair, as Devitt and Porot fail to appropriately shed light on the different formulations of DT, which comprises both a weak version and a strong one. Moreover, while Devitt and Porot have the merit of emphasizing the varying descriptions that different groups of speakers can associate with a single name, thereby determining – according to DT – different referents, Devitt and Porot's theoretical

treatment of what they term the "New Meaning Objection" is not totally satisfactory.[145] First, the issue potentially afflicting their experiment on proper names is not the abrupt introduction of a potentially new reference-fixing description. Rather, the predicament is that the descriptivist may take the narrator to rely upon a previously established reference-fixing description being different from the one in force within the community that the vignette presents. Devitt and Porot's vignette fails to address that problem, contrary to their claim. However, I finally argue that the way they experimentally tackle the New Meaning Objection in their follow-up study is elegant and compelling. Therefore, Devitt and Porot ultimately provide experimental data that support CHT for their American sample.

As said, Devitt and Porot (2018) pertains to linguistic production. However, it is crucial to also explore the complementary aspect of linguistic usage, which involves comprehension. The corroboration of CHT is more robust if the data stemming from that side of usage too converge in the same direction. Domaneschi and Vignolo (2020) and Li (2021) have the merit of trying to test linguistic usage while considering the potentially confounding epistemic ambiguity. Furthermore, although the authors fail to emphasize this aspect, their studies indirectly control for the potential attributive reading of names. However, the way Domaneschi and Vignolo frame the epistemic ambiguity could be misleading, since it might suggest that the ambivalence at stake is between two equally valid meanings of the truth predicate, while it actually depends on whether participants understand the predicate literally or not. Indeed, the very term "ambiguity", as applied to the epistemic ambivalence at stake, is imprecise – during this dissertation I have used it just as a convenient shorthand. Moreover, the way Domaneschi and Vignolo try to experimentally control for the epistemic confounding factor is disputable, both for the phrasing of the follow-up options, which is biased against DT, and for the choice of using a follow-up question in the first place, which reignites Martí's criticism. Finally, Domaneschi and Vignolo's vignette suffers from the potential problem of the social interpretation of predicates. In light of these reasons, I question the Kripkean conclusion that Domaneschi and Vignolo draw from their data.

[145] If different groups of speakers associate different descriptions with a name, then, according to DT, different referents ensue and arguably that name is no longer a single name; rather, there are different (descriptive) names. Therefore, strictly speaking, those groups of speakers would be associating different descriptions not with a single name but with a single orthographic type.

Li (2021) is structurally similar to Domaneschi and Vignolo (2020). Li too aims to control for the epistemic ambiguity with a follow-up question. As her follow-up question has an open format, she avoids the problem of the possibly biased formulation of the follow-up answers, which afflicts Domaneschi and Vignolo's experiment. Moreover, she wisely uses predicates that sidestep a potential social interpretation. Li concludes that the cross-cultural difference in the TVJs that she collects does not reflect a variation between DT and CHT, but rather stems from culturally specific strategies of perspective-taking. I agree that Li's data do not prove a cross-cultural difference in the way people use names, but I question the "perspectives" that Li targets. She claims that the cross-cultural difference emerging in her Original Condition stems from the possibility to interpret the final prompt as conceptually analogous either to the one of the Internal Condition or to one of the External Condition. Unfortunately, the latter fails to capture the approach of the participants evaluating the factual truth value in the Original Condition, as the External Condition introduces a speaker – Kermit – who is not in semantic contradiction. However, the External Condition indirectly aids in detecting the specific non-literal interpretation that might guide the participants not judging the factual truth value of the critical sentence. That is a non-literal reading involving the perspective not of the single speaker (e.g., Emily), but rather of the broader community to which she belongs. As regards the Original Condition, it fails to control for the epistemic ambiguity because it uses a follow-up question, which reignites Martí's criticism, analogously to what happens in Domaneschi and Vignolo (2020). For these reasons, I also reject Li's (2023a) subsequent reinterpretation of her results, according to which her data support a pluralist position. I finally argue that the remarkably high proportion of TVJs "False" that Li collects with her American sample corroborates the thesis that the attributive reading of names is a confounding factor that has indeed afflicted the previous experimental literature – as Heck (2018) and Vignolo and Domaneschi (2022) indicate. That said, in light of a comprehensive analysis of the literature with American and Western participants (including also, for instance, Study 3 in Li, 2023a), it is likely hasty to conclude that Li's results definitely confirm CHT as the correct theory of reference for that demographic. Further research is needed, and Chapter 3 and 4 try to address this gap.

In Chapter 3, I present my Experiment 1 and 2, which draw methodological inspiration from Islam and Baggio (2020) and try to control for the epistemic ambiguity without

relying upon follow-up questions at all (Experiment 2) or ascribing them only a secondary role (Experiment 1). Experiment 1, conducted in collaboration with Litman Huang, presents some explorative TVJ data on the Jonah Case with an Italian sample and a Chinese one. The tentative nature of the data of Experiment 1 is due to its partially flawed design. Despite these limitations, the TVJ part of the experiment is sufficient to address the main experimental question and corroborate CHT as the correct theory for both samples: only a limited number of the Italian and Chinese participants opt for the TVJ "False", which is a significant finding because in the Jonah Case that TVJ can only express a descriptivist understanding of the name. Therefore, the cross-cultural variation with the Gödel Case that the literature unveiled over the years seems not to have a semantic root but rather derives from the different perspective-taking strategies across various cultural backgrounds. Experiment 2 involves only Italian speakers and presents both Gödel vignettes and Jonah ones to the participants. In this way, it is possible to ascertain how the single participants behave across the two scenarios, and thus establish whether the TVJ "True" in the Gödel Case stems from a genuine descriptivist understanding of names or rather from a non-literal interpretation of the truth predicate. The experiment, which also controls for the social interpretation of predicates and the attributive understanding of names, corroborates CHT as the correct theory for my Italian sample: the Gödel scenarios elicit a limited proportion of supposedly descriptivist TVJs, and the Jonah scenarios show that those arguably derive from the adoption of the community's perspective.

In Chapter 4, I present Experiment 3, an eye-tracking study that Filippo Domaneschi, Massimiliano Vignolo, Camilo Rodríguez Ronderos and I conduced with a sample of Italian participants. The rationale of the experiment is to collect more direct data concerning linguistic comprehension. Following some suggestions that Cohnitz and Haukioja (2015) and Cohnitz (2015) put forth, we find a suitable technique in the eye-tracking methodology, which has a long-established tradition in psycholinguistics. While TVJs are supposed to unveil the proposition that the participant understands as expressed by a certain sentence after the comprehension process, the eye-tracking methodology sheds light on the referential connections that the participant establishes while her understanding process unfolds. Our experiment locates itself in the context of the previous eye-tracking literature that shows that people's eye movements are sensitive to

the mental contents that the vignette ascribes to the characters within the story. That is a crucial aspect, as any Gödel-style scenario relies upon John's state of semantic contradiction: the participant must be sensitive to the fact that John's description picks out an individual different from the one causally-historically connected to the name. Without that sensitivity, her answer is unsuitable to compare the two theories. Our eye-tracking results corroborate CHT as the correct theory of reference for our Italian sample. Therefore, at first glance, these findings seem to indirectly support the reliability of TVJs, especially in light of Experiment 2's validation of the previous tendency of the Gödel Case to elicit remarkably high proportions of causal-historical TVJs among Westerners when the experimenter controls for the social interpretation of the predicate and the attributive understanding of names (Li, 2021). The limited residual percentage of purportedly descriptivist TVJs likely stems from the adoption of the perspective of the speaker's community. However, the evaluation of the reliability of TVJs becomes more complicated if one considers the TVJ results that we collected *within* our eye-tracking experiment. They strongly align with the descriptivist prediction ("True"), thereby substantiating Martí's (2012, 2014, 2020) reservations about the reliability of TVJs when the test involves a speaker like John or Emily. That said, the extensive deviation of these TVJs from the previous literature may invite this interpretative hypothesis: TVJs are generally reliable, and the unusual TVJ results in the context of our experiment may be due to some confounding variables that specifically stem from our innovative eye-tracking design and still lack precise identification. The experimental setting in which the TVJ prompt appears may somehow prompt participants to formulate their TVJ by adopting the community's perspective or engaging in the kind of theoretical reflection that Martí stresses. Further research is required on this topic.

At this point, one may note that the good standing of CHT that this thesis defends was also – so to speak – the starting point before the experimental debate that MMNS initiated, since the philosophical community generally accepted Kripke's refutation of descriptivism. Therefore, one could object that the outcome that experimental methods reached merely reaffirms a conclusion that the philosophical community had already embraced. In other words, the experimental debate would fail to offer substantially new insights. Needless to say, I disagree with that thesis. What matters in philosophy is not only the conclusion that one reaches. What is also of paramount importance are the

arguments that lead to a specific conclusion. Kripke's causal-historical view is a descriptive enterprise that aims to account for how speakers in fact use names. Speakers' usage is an empirical matter that scholars cannot settle *a priori*. Therefore, absent a specific empirical investigation aimed to ascertain how speakers in fact use names in situations like the Gödel one, any conclusion as to the correct theory of reference relies upon conjectural grounds. This dissertation shows that, at least with Italian speakers, philosophers' referential intuitions in the Gödel Case endure the scrutiny that the empirical challenge poses.

A crucial point concerns the demographic generalizability of my conclusions. MMNS's main focus is precisely that laypeople's referential intuitions may vary across countries and cultural backgrounds – as well as within those. If one accepts MMNS's (2004) comparison between "Westerners" and "East Asians", one might try to extend my Kripkean conclusions from Italian speakers to American ones.[146] That extension may seem warranted also in light of Devitt and Porot's (2018) follow-up experiment, which successfully deals with the New Meaning Objection and provides linguistic-production evidence in support of CHT with American participants. However, as said, MMNS (2004) has a crucial cross-cultural focus as well. As regards that, my Experiment 1, despite providing admittedly only provisional data, suggests that also Chinese's linguistic usage aligns with CHT. Therefore, one may conjecture that MMNS's cross-cultural finding does not have a semantic nature, that is, the variation between Chinese and people from other countries such as Italy or USA is not rooted in different ways of using names. It is also possible to advance some other, convergent considerations in support of that thesis. As seen, the limited seemingly descriptive evidence stemming from my Gödel scenarios with Italian participants in Experiment 2 does not reflect a truly descriptivist stance. Interestingly, those results might indirectly corroborate the overall reliability of the follow-up explanations that the Italian sample in Domaneschi and Vignolo (2020) provides. When formulating a follow-up answer to explain the TVJ "True" in the Gödel Case, Domaneschi and Vignolo's participants appeal to the speaker's perspective and not

[146] Note however that MMNS are aware of the potential oversimplification in using the terms "Westerners" and "East Asians": «There is a common concern that the labels 'East Asian' and 'Western' are too rough to do justice to the enormous diversity of cultural groups such labels encompass. We are sympathetic to this concern. However, the crudeness of these groupings does nothing to undermine the experiment we present. On the contrary, if we find significant results using crude cultural groupings, there is reason to believe more nuanced classifications should yield even stronger results» (MMNS, 2004, p. B5, n. 9).

to a descriptivist rationale. Those explanations square with my results, which provide evidence against DT and in favor of CHT. In other words, Italian participants are not descriptivists and indeed, when justifying the TVJ "True", they reliably provide a non-descriptivist explanation.[147] Crucially, in Li (2021) Chinese participants also justify the TVJ "True" by appealing to the community's perspective. If one inductively extends the reliability of the follow-up answers to that Chinese sample, one may conclude that among Chinese too the supposedly descriptivist evidence is not genuinely such. As a consequence, the cross-cultural variation has no semantic root. More in general, in light of the confounding factors whose influence my experiments highlight, the collected data at least suggest that it would be too hasty and naïve to conclude that Chinese speakers are descriptivist users of names. That said, future more direct research, such as in the domain of eye-tracking, will obviously provide more solid evidence as to how Chinese and East Asians use names.

A natural question that arises at this point regards what semantic fields, beyond proper names, will be at the center of future experimental research on reference. Several scholars have already investigated natural-kind terms experimentally, and the reason is plain to see: that is the other category of terms for which Kripke proposes CHT.[148] Another semantic area concerns artifact terms, to which Putnam (1975) extends the casual-historical approach: to date, the only published study in the field of experimental philosophy investigating (also) that category of words is Genone and Lombrozo (2012).

More in general, irrespective of the semantic field of application, testing a theory of reference is also an empirical enterprise: that is the fundamental and invaluable lesson that derives from MMNS's seminal work. The other crucial point that emerges from the debate subsequent to MMNS (2004) is that not all the kinds of data are equally relevant: scholars' main focus has to be on linguistic facts rather than on laypeople's meta-linguistic reflections about them, as there is no *a priori* guarantee that those reflections

[147] As seen, Domaneschi and Vignolo's work suffers from the independent problem of proposing a descriptivist explanation that incorporates an implicit bias against DT (a bias that a descriptivist participant may overcome by ignoring the remark "unbeknown to Tom and Emily's teacher"). Note, however, that Devitt and Porot's (2018) linguistic-production data also provide evidence in support of CHT with an American sample and, consistently with that, the few American participants who in Li's (2021) Experiment 1 and in Li's (2023a) Study 3 choose the TVJ "True" do not provide a descriptivist explanation.

[148] See Braisby et al. (1996), Jylkkä et al. (2009), Genone and Lombrozo (2012), Buckwalter and Stich (2014), Nichols et al. (2016), Tobia et al. (2020), Devitt and Porter (2021), Haukioja et al. (2021), Koch and Wiegmann (2022), Devitt and Porter (2023), Haukioja et al. (2023), Machery et al. (2023). For discussion on the studies, see Hansen (2015), Martí (2015), Cohnitz and Haukioja (2020), Dacey and Mallon (2016), and Machery (2023).

are reliable. What philosophers need to do is to temporarily step away from their armchairs to collect the most direct linguistic data, rather than «pulling out more armchairs for people to sit on» (Martí, 2020, p. 331)[149]. At the end of the data-collection stage, philosophers should then sit again on their armchairs to reflect on what their findings mean: data «don't interpret themselves» (Appiah, 2007) and, without the proper philosophical speculation, any empirical result remains silent.

[149] Martí (2020, p. 331, n. 4) writes: «for the longest time, I was sure I owed that humorous remark to Michael Devitt. But he assures me that he actually heard it from me». Whoever the originator of the remark is, I thank her/him because it brilliantly summarizes the approach that experimental philosophers should avoid. Naturally, that does not mean that laypeople's meta-linguistic reflections are always inaccurate: while my studies indicate that laypeople's RIs do not align with how speakers use proper names, I suggested that the participants' justificatory remarks for their TVJs in Domaneschi and Vignolo (2020) and Li (2021) conform to how they in fact use names. However, that is a conclusion that one can reach only *a posteriori*, by comparing the meta-linguistic reflections with linguistic usage, namely, with the reality that theories of reference seek to explain.

Appendix

Experiment 1 involves two phases: a TVJ one and a follow-up one, the latter delving into the explanations that the participant provides to justify her previous TVJ. However, as explained (section 3.2), the follow-up part – on top of being theoretical – suffers from some wording shortcomings. In this appendix, I present the follow-up material and the related results, and I explain the limitations in the formulation of the options.

An instance of TVJ final prompt of Experiment 1 is this:

> Leonardo is a history student who carefully attended the lessons on Roman invasions. Like any other person, the only description that Leonardo associates with the name "Ambiorix" is "The warrior who conducted the Gaul army when the Romans attacked Gaul".
>
> When Leonardo says «Ambiorix existed», what Leonardo says is
> - True
> - False
> - I don't know

As seen (section 3.2.4), across the six targets, Italian participants select the TVJ "True" 90.93% of the time, the TVJ "False" 4.66% and the "I don't know option" 4.41%. Chinese participants choose the TVJ "True" 81.48% of the time, the TVJ "False" 7.41% and the "I don't know option" 11.11%.

If the participant chooses the TVJ "True", she is presented with this follow-up question:

> You answered "true" because
>
> (a) Leonardo is talking about the peacetime leader, although Leonardo erroneously believes that that person was a warrior who conducted the Gaul army when the Romans attacked Gaul.
> (b) Leonardo reports what he was told, namely, that Ambiorix existed.
> (c) Leonardo means: "Ambiorix, whoever he was, existed".

Option (a) is supposed to capture the TVJ of the participant who chooses "True" for genuinely causal-historical reasons: Leonardo is talking about the individual at the end of the causal-historical chain, Leonardo's belief is about that individual and it is wrong. Option (b) is supposed to capture the TVJ of the participant who answers "True" not because she is concerned with the factual truth value of the sentence, but rather because she wants to stress that Leonardo's statement aligns with the consensus within his community, that is, Leonardo correctly repeats what he has been told. Option (c), in theory, is supposed to capture the answer of that participant who imagines Leonardo as claiming that Ambiorix, regardless of whether he conducted the Gauls against the Romans, existed. Compared to (a), option (c) delineates a situation in which Leonardo temporarily sets aside the description that he associates with the name "Ambiorix". We conceived (c) as an option accessible only to the causal-historical participant, since a descriptivist participant could not imagine a situation in which Leonardo disregards that description. According to DT, that description is what mediates the name's reference, and thus Leonardo cannot use the name while supposing away its descriptive basis. However, the formulation of (c) is unfortunate. One matter is to claim that Leonardo means "Ambiorix, irrespective of whether he conducted the Gauls against the Romans, existed": this formulation explicitly specifies the description that Leonardo puts aside while using the name. Another matter is to claim, as (c) does, that Leonardo means: "Ambiorix, whoever he was, existed": this formulation is way vaguer and, crucially, is consistent also with DT. Consider a prototypical descriptive name like "Jack the Ripper": one can assert "Jack the Ripper, whoever he was, existed", meaning that the killer of a certain group of prostitutes in 1888 existed, whoever that individual exactly was (say, a middle-age man as opposed to an old one, or a very rich person as opposed to a poor one). In other words, the statement "X, whoever he/she was, existed" approximates a mere rephrasing of the critical statement "X existed". Thus, option (c) fails to reveal how participants justify their TVJ on Leonardo's utterance, thereby partially compromising the analysis of the follow-up results.

Figure 1 Appendix reports the follow-up results concerning the TVJ "True".

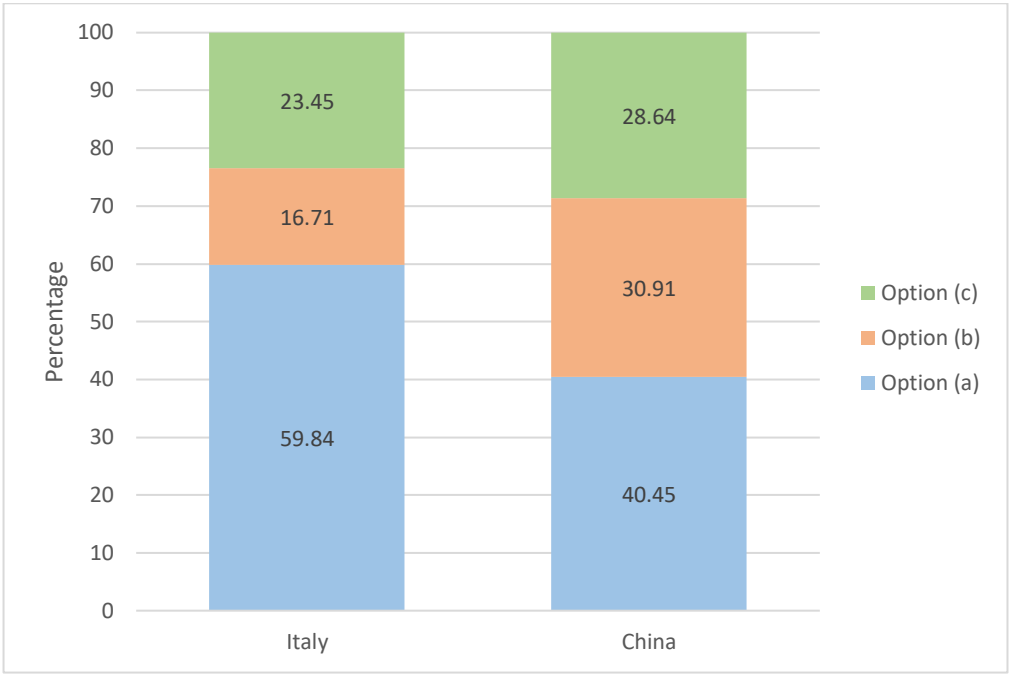

Figure 1 Appendix Percentages of explanations after "True" in Experiment 1

Within the Italian sample, the Chi-square statistics confirm that the difference in frequency between (a) and both (b) and (c) (respectively: χ^2 (1, N = 284) = 90.14, p < .001; χ^2 (1, N = 309) = 58.98, p < .001), and between (b) and (c) (χ^2 (1, N = 149) = 4.20, p < .05). Within the Chinese sample, the difference between (a) and (c) is significant, χ^2 (1, N = 152) = 4.48 p < .05), but not the difference between (a) and (b), and (b) and (c) (ps > .05). To determine whether the explanations' frequencies differ across cultures, a multinomial logistic mixed effects model with culture as fixed effect and participant as random effect was constructed. The model reveals a significant effect of culture (F (2, 587) = 5.59, p < .01). The fixed coefficients of the pairwise comparisons of the multinomial model have option (a) as the baseline category. The comparisons reveal that that Italians are significantly less inclined than Chinese to choose option (b) than option (a) (ß = -1.07, SE = .34, t = -3.16, p < .01, Exp(ß) = .34), while no statistical difference between (c) and (a) emerges (ß = -.66, SE = .44, t = -1.50, p >.05, Exp(ß) = .52).[150]

[150] In the initial phase of our analysis, the model also included the random intercept for vignette and the slope of culture random across vignettes, but these random effects were removed. Their variance in the pairwise comparison between option (a) and (b) was approximately 0, leading to a non-positive-definite final Hessian matrix of the multinomial model, which may make the validity of the model fit uncertain. That said, the statistics of the model including both the

Table 1 Appendix and Table 2 Appendix report the percentages across the six targets for the Italian and Chinese sample.

Table 1 Appendix Percentages of explanations after "True" for the six items in Experiment 1 – Italian sample

	Ambiorix Case	Anaximenes Case	Collaborator Case	Fabrication Case	Murder Case	Sister Case
Option (a)	55.93	54.24	61.90	63.08	58.73	64.52
Option (b)	15.25	16.95	14.29	16.92	20.63	16.13
Option (c)	28.81	28.81	23.81	20.00	20.63	19.35

Table 2 Appendix Percentages of explanations after "True" for the six items in Experiment 1 – Chinese sample

	Ambiorix Case	Anaximenes Case	Collaborator Case	Fabrication Case	Murder Case	Sister Case
Option (a)	23.68	37.14	57.89	47.22	40.54	36.11
Option (b)	36.84	28.57	31.58	27.78	37.84	22.22
Option (c)	39.47	34.29	10.53	25.00	21.62	41.67

While the mixed model, which combines the six items, reveals a significant effect of culture, the analysis of the single items reveals a significant effect of culture only in the Ambiorix Case (χ^2 (2, N = 97) = 10.89, $p < .01$; Cramer's V = .34) and the Sister Case (χ^2 (2, N = 98) = 7.97, $p < .05$; Cramer's V = .29), despite the wording's structural similarity across all the items.[151] Nevertheless, some constant tendency emerges from the descriptive statistics for the single items: the percentages for option (a) and (b) in the Italian sample are always numerically higher and lower, respectively, compared to those in the Chinese sample. Moreover, except for the Collaborator Case, the percentages for option (c) are always numerically lower in the Italian sample than in the Chinese one. Possibly, a significant effect emerges only in two of the six items because of the limited number of observations in 2×3 contingency tables – a larger sample might have revealed

slope of culture and the random intercept for vignette are virtually the same as the one not including it. Omnibus test: F (2, 587) = 5.40, $p < .01$; option (a) vs. option (b): ß = -1.07, SE = .34, t = -3.16, $p < .01$, Exp(ß) = .34; option (a) vs. option (c): ß = -.64, SE = .50, t = -1.27, $p > .05$, Exp(ß) = .53).

[151] The reported values for Cramer's V correspond to a moderate effect size (IBM website, Cramér's V, <https://www.ibm.com/docs/en/cognos-analytics/12.0.0?topic=terms-cramrs-v>, consulted on December 18th, 2023). Note that also the Collaborator Case approximates significance (χ^2 (2, N = 101) = 5.70, $p = .058$; Cramer's V = .24).

significant effects for the other items as well. Finally, as seen, the pairwise comparisons of the mixed models signal that Chinese are comparatively more inclined than Italians to choose option (b) as opposed to option (a). Therefore, in a speculative way, one may advance the following conjecture. As seen, option (a) stresses the Leonardo's belief is erroneous. In accordance with Devitt's (2023) hypothesis, one may suppose that Chinese are less inclined than Italians to choose (a) because people from an East Asian background may have more difficulties to claim that one is wrong if one is blamelessly such. As a consequence, participants from a Chinese background might be comparatively more inclined to select the alternative choices. However, this conclusion is speculative given the non-informative option (c), which taints the analysis of the results.

We presented a follow-up question also to those participants who select the TVJ "False".

> You answered "false" because
>
> (d) No one existed who has the features that Leonardo associates with the name "Ambiorix", namely being a warrior who conducted the Gaul army against the Romans
> (e) A person named "Ambiorix" did not exist

Option (d) is supposed to capture the descriptivist stance: Leonardo's claim is false because what mediates the name's reference is the description and no entity satisfies it. On the other hand, option (e) was initially intended to serve as a further attention check alongside the filler questions: since the vignette uses the name "Ambiorix" to talk about an individual, there must have been an individual known by that name in the past. Thus a participant should not choose an option to the effect that no one bearing that name existed. However, upon further reflection, we realized that (e) is compatible with a descriptivist point of view: a descriptivist can choose (e) to stress that, within Leonardo's language, that name is empty and thus a person bearing it never existed.

Within both groups, participants opt for the TVJ "False" quite rarely. As regards the Italian sample, out of 68 participants and over the course of the six vignettes, they choose the TVJ "False" only 19 times (for a total of 4.66% of the entire group of TVJs). As regards the Chinese sample, out of 45 participants and over the course of the six vignettes,

they choose the TVJ "False" 20 times (for a total of 7.41% of the entire group of TVJs). In both cases, participants mostly select the follow-up option (d): the Italian sample 78.95% of the time, the Chinese sample 90% of the time. A binary logistic mixed effects model with culture as fixed effect and participant as random effect was constructed to determine whether there are any cultural differences in the explanation pattern across cultures. The model reveals no significant effect of culture (ß = -.83, SE = 1.07, t = -.78, $p > .05$, Exp(ß) = .43).[152] Arguably, participants opt for (d) over (e) because the former expresses a descriptivist stance in a more direct and straightforward way. However, as the participants opting for the TVJ "False" in the first place are a strict minority, these supposedly descriptivist answers are arguably philosophically inconsequential.[153]

[152] Initially, the model also included the random intercept for vignette and the slope of culture random across vignettes, but these random effects were removed because they led to a non-positive-definite final Hessian matrix. Their variance was 0, and indeed statistics of the complete model are exactly the same as the one not including them.

[153] For the sake of completeness, this note details the follow-up results of all the participants taking the Italian and the Chinese versions of the questionnaire, including those not meeting the inclusion criteria. As regards the follow-up question after the TVJ "True", the percentages of explanation within the Italian sample are these: option (a) 58.74% of the time, (b) 19.37%, (c) 21.89%. The frequency of (a) differs statistically significantly from (b) (χ^2 (1, N = 371) = 94.26, $p < .001$) and (c) (χ^2 (1, N = 383) = 79.96, $p < .001$), while (b) does not differ from (c) ($p > .05$). Within the Chinese sample, the percentages are these: (a) 38.15% of the time, (b) 27.18%, (c) 34.66%. The frequency of (a) differs statistically significantly from (b) (χ^2 (1, N = 262) = 7.39, $p < .01$), but not from (c) ($p > .05$); (b) approximates significance in comparison with (c) (χ^2 (1, N = 248) = 3.63, $p = .057$). A multinomial logistic mixed effects model, with culture as fixed effect and participant as random effect, reveals that the cultural background has a significant impact on the explanation (F (2, 872) = 7.31, $p < .001$). The fixed coefficients of the pairwise comparisons of the multinomial model have option (a) as the baseline category. The comparisons reveal that that Italians are significantly less inclined compared to Chinese to choose option (b) (ß = -.83, SE = .31, t = -2.70, $p < .01$, Exp(ß) = .44) and option (c) (ß = -1.08, SE = .36, t = -3.01, $p < .01$, Exp(ß) = .34) than option (a). (The model omits the random intercept for vignette and the random slope for culture across vignettes because otherwise the model would have had a non-positive-definite Hessian matrix; anyhow, a model including those random effects leads to the same significances and the same levels of significance.) Table Appendix 3 reports the percentages for each target in both the Italian and Chinese sample.

Table 3 Appendix Percentages of explanations after "True" for the six items in Experiment 1

■ Italian sample ■ Chinese sample

	Ambiorix Case		Anaximenes Case		Collaborator Case		Fabrication Case		Murder Case		Sister Case	
Option (a)	55.84	29.41	54.05	39.68	62.50	50.70	60.24	38.24	57.32	35.82	62.03	34.38
Option (b)	19.48	35.29	20.27	23.81	15.00	26.76	19.28	26.47	23.17	29.85	18.99	20.31
Option (c)	24.68	35.29	25.68	36.51	22.50	22.54	20.48	35.29	19.51	34.33	18.99	45.31

The analysis of the items reveals a significant effect of culture in the Ambiorix Case (χ^2 (2, N = 145) = 10.54, $p < .01$; Cramer's V = .27), the Fabrication Case (χ^2 (2, N = 151) = 7.48, $p < .05$; Cramer's V = .22), the Murder Case (χ^2 (2, N = 149) = 7.30, $p < .05$; Cramer's V = .22) and the Sister Case (χ^2 (2, N = 143) = 13.44 $p < .001$; Cramer's V = .31). As regards the Italian sample, out of 91 participants and over the course of six vignettes, they choose the TVJ "False" only 27 times, accounting for 4.95% of all TVJs. As regards the Chinese sample, out of 84 participants and over the course of six vignettes, they chose the TVJ "False" 40 times, comprising for 7.94% of all TVJs. In both cases, participants mostly select the follow-up option (d): the Italian sample 74.07% of the time, the Chinese sample 90%. A binary logistic mixed effects model, with culture as fixed effect and participant and vignette as random effects, shows that the cultural background does not influence the choice of the explanation (ß = -.76, SE = .68, t = -1.11, $p > .05$, Exp(ß) =.47). (This binary mixed effect model, unlike the one of the main analysis, includes the random effect of vignette because its variance is different from 0. The model omits the random slope for culture across vignettes because otherwise the model would have had a non-positive-definite Hessian matrix; anyhow, a model including that random effect leads to the same non-significant result.) Thus, the results obtained from the entire group of survey respondents are broadly consistent with those from the sample included in the final analysis.

References

Achenbach J (2005) Cat Carrier: Your cat could make you crazy. *National Geographic* 208.

Adams Z, Hansen N (2020) The myth of the common-sense conception of colour. In *Shifting Concepts. The Philosophy and Psychology of Conceptual Variability*, Marques T, Wikforss Å (eds) pp. 106-127. New York: Oxford University Press. https://doi.org/10.1093/oso/9780198803331.003.0007

Alexander J, Mallon R, Weinberg JM (2010) Accentuate the negative. *Review of Philosophy and Psychology* 1(2): 297-314. https://doi.org/10.1007/s13164-009-0015-2

Altmann GTM, Kamide Y (1999) Incremental Interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73(3): 247-264. https://doi.org/10.1016/S0010-0277(99)00059-1

Appiah KA (2007) The New New Philosophy. *The New York Times Magazine*, 9 December, <https://www.nytimes.com/2007/12/09/magazine/09wwln-idealab-t.html>.

Arai M, Van Gompel RP, Scheepers C (2007) Priming ditransitive structures in comprehension. *Cognitive Psychology* 54(3): 218-250. https://doi.org/10.1016/j.cogpsych.2006.07.001

Barnard R, Ulatowski J (2021) The objectivity of truth, a core truism?. *Synthese* 198(Suppl 2): 717-733. https://doi.org/10.1007/s11229-017-1605-7

Beebe JR (2024) Experimental epistemology: knowledge and Gettier cases. In *The Compact Compendium of Experimental Philosophy*, Bauer AM, Kornmesser S (eds) pp. 163-182. Berlin and Boston: Walter de Gruyter GmbH. https://doi.org/10.1515/9783110716931-008

Beebe JR, Undercoffer RJ (2015) Moral valence and semantic intuitions. *Erkenntnis* 80(2): 445-466. https://doi.org/10.1007/s10670-014-9653-6

Beebe JR, Undercoffer R (2016) Individual and cross-cultural differences in semantic intuitions: New experimental findings. *Journal of Cognition and Culture* 16(3/4): 322-357. https://doi.org/10.1163/15685373-12342182

Braisby N, Franks B, Hampton J (1996) Essentialism, word use, and concepts. *Cognition* 59(3): 247-274. https://doi.org/10.1016/0010-0277(95)00698-2

Buckwalter W (2014) Factive verbs and protagonist projection. *Episteme* 11(4): 391-409. https://doi.org/10.1017/epi.2014.22

Buckwalter W (2022) The replication crisis and philosophy. *Philosophy and the Mind Sciences*, 3. https://doi.org/10.33735/phimisci.2022.9193

Buckwalter W, Stich S (2014) Gender and philosophical intuition. In *Experimental Philosophy*. Vol. II, Knobe J, Nichols S (eds) pp. 307-346. New York: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199927418.003.0013

Burge T (1979) Individualism and the mental. *Midwest Studies in Philosophy* 4(1): 73-121. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Cappelen H (2012) *Philosophy without Intuitions*. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780199644865.001.0001

Carston R (2002) *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Blackwell, Oxford. https://doi.org/10.1002/9780470754603

Choi I, Nisbett RE, Norenzayan A (1999) Causal attribution across cultures: Variation and universality. *Psychological Bulletin* 125(1): 47-63. https://doi.org/10.1037/0033-2909.125.1.47

Cohnitz D (2015) The metaphilosophy of language. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 85-108. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-005

Cohnitz D, Häggqvist S (2009) The role of intuitions in philosophy. *Studia Philosophica Estonica* 2(2): 1-14. https://doi.org/10.12697/spe.2009.2.2.01

Cohnitz D, Haukioja J (2013) Meta-externalism vs meta-internalism in the study of reference. *Australasian Journal of Philosophy* 91(3): 475-500. https://doi.org/10.1080/00048402.2012.717534

Cohnitz D, Haukioja J (2015) Intuitions in philosophical semantics. *Erkenntnis* 80(3): 617-641. https://doi.org/10.1007/s10670-014-9666-1

Cohnitz D, Haukioja J (2020) Variation in natural kind concepts. In *Shifting Concepts. The Philosophy and Psychology of Conceptual Variability*, Marques T, Wikforss Å (eds) pp. 551-559. New York: Oxford University Press. https://doi.org/10.1093/oso/9780198803331.003.0008

Cohnitz D, Haukioja J (2021) Reference and intuitions. In *The Routledge Handbook of Linguistic Reference*, Biggs S, Geirsson H (eds) pp. 551-559. New York and Abingdon: Routledge. https://doi.org/10.4324/9781003111894-51

Cova F, Strickland B, Abatista A, Allard A, Andow J, Attie M, Beebe J, Berniūnas R, Boudesseul J, Colombo M, et al. (2018) Estimating the reproducibility of experimental philosophy. *Review of Philosophy and Psychology* 12(1): 9-44. https://doi.org/10.1007/s13164-018-0400-9

Cummins R (1998). Reflection on reflective equilibrium. In *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*, DePaul M, Ramsey W (eds) pp. 113-127. Oxford: Rowman & Littlefield.

D'Agruma N (2022) Riferimento dei nomi propri e filosofia sperimentale. *APhEx* 25.

D'Agruma N (2023) From the epistemic perspectives in experimental semantics to the ambiguity of proper names: Is the inference warranted? A critical notice of Jincai Li's *The referential mechanism of proper names*. *Mind & Language* 38(4): 1138-1146. https://doi.org/10.1111/mila.12472

D'Agruma N (manuscript) Social facts, circularity and causal-historical connections in experimental semantics: where Ding and Liu go wrong, under review.

Dacey M, Mallon R (2016) Reference. In *A Companion to Experimental Philosophy*, Sytsma J, Buckwalter W (eds) pp. 371-389. Oxford: Blackwell. https://doi.org/10.1002/9781118661666.ch26

Deutsch M (2009) Experimental philosophy and the theory of reference. *Mind & Language* 24(4): 445-466. https://doi.org/10.1111/j.1468-0017.2009.01370.x

Deutsch M (2010) Intuitions, counter-examples, and experimental philosophy. *Review of Philosophy and Psychology* 1: 447-460. https://doi.org/10.1007/s13164-010-0033-0

Deutsch M (2015) Kripke's Gödel Case. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 7-30. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-002

Devitt M (1981) *Designation*. Columbia University Press, New York.

Devitt M (1984/1997) *Realism and Truth*. Princeton University Press, Princeton.

Devitt M (2011) Experimental semantics. *Philosophy and Phenomenological Research* 82(2): 418-435. https://doi.org/10.1111/j.1933-1592.2010.00413.x

Devitt M (2012a) Semantic epistemology: Response to Machery. *THEORIA. An International Journal for Theory, History and Foundations of Science* 27(2): 229-233. https://doi.org/10.1387/theoria.6225

Devitt M (2012b) Whither experimental semantics?. *THEORIA. An International Journal for Theory, History and Foundations of Science* 27(1): 5-36. https://doi.org/10.1387/theoria.3866

Devitt M (2012c) The role of intuitions. In *The Routledge Companion to Philosophy of Language*, pp. 554-565. New York and Abingdon: Routledge. https://doi.org/10.4324/9780203206966.ch4_6

Devitt M (2015a) Should proper names still seem so problematic? In *On Reference*, Bianchi A (ed) pp. 108-143. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198714088.003.0007

Devitt M (2015b) Testing theories of reference. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 157-172. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-003

Devitt M (2015c) Relying on intuitions: Where Cappelen and Deutsch go wrong. *Inquiry* 58(7-8): 669-699. https://doi.org/10.1080/0020174X.2015.1084824

Devitt M (2020) Stirring the possum: Responses to the Bianchi papers. In *Language and Reality from a Naturalistic Perspective: Themes from Michael Devitt*, Bianchi A (ed) pp. 371-455. Cham: Springer. https://doi.org/10.1007/978-3-030-47641-0_19

Devitt M (2021) *Overlooking Conventions: The Trouble with Linguistic Pragmatism*. Springer, Cham. https://doi.org/10.1007/978-3-030-70653-1

Devitt M (2023) Red herrings in experimental semantics: Cultural variation and epistemic perspectives. A critical notice of Jincai Li's *The referential mechanism of proper names*. *Mind & Language* 38(4): 1147-1156. https://doi.org/10.1111/mila.12468

Devitt M, Porot N (2018) The reference of proper names: Testing usage and intuitions. *Cognitive Science* 42(5): 1552-1585. https://doi.org/10.1111/cogs.12609

Devitt M, Porter BC (2021) Testing the reference of biological kind terms. *Cognitive Science* 45(5): e12979. https://doi.org/10.1111/cogs.12979

Devitt M, Porter BC (2023) Two sorts of biological kind terms: The cases of 'rice' and 'Rio de janeiro Myrtle'. *Philosophy and Phenomenological Research* 1-27. https://doi.org/10.1111/phpr.12979

Devitt M, Sterelny K (1999) *Language and Reality: An Introduction to the Philosophy of Language*. Blackwell Publishers, Oxford.

Ding C, Liu C (2022) Kripke's Gödel case: Descriptive ambiguity and its experimental interpretation. *THEORIA. An International Journal for Theory, History and Foundations of Science* 37(3): 291-308. https://doi.org/10.1387/theoria.23375

Domaneschi F, D'Agruma N, Vignolo M, Ronderos CR (manuscript) Reference and proper names. A clash between eye-tracking and truth-value judgments, under review.

Domaneschi F, Di Paola S (2019) Relevance and non-factive knowledge attributions. *Acta Analytica* 34(1): 83-115. https://doi.org/10.1007/s12136-018-0362-7

Domaneschi F, Vignolo M (2020) Reference and the ambiguity of truth-value judgments. *Mind & Language* 35(4): 440-455. https://doi.org/10.1111/mila.12254

Domaneschi F, Vignolo M, Di Paola S (2017) Testing the causal theory of reference. *Cognition* 161: 1-9. https://doi.org/10.1016/j.cognition.2016.12.014

Donnellan K (1966) Reference and definite descriptions. *Philosophical Review* 75(3): 281-304. https://doi.org/10.2307/2183143

Donnellan K (1970) Proper names and identifying descriptions. *Synthese* 21(3/4): 335-358. https://doi.org/10.1007/BF00484804

Dummett M (1973) *Frege: Philosophy of Language*. Duckworth, London.

Evans G (1973) The causal theory of names. *Proceedings of the Aristotelian Society, Supplementary Volumes* 47(1): 187-208. https://doi.org/10.1093/aristoteliansupp/47.1.187

Evans G (1979) Reference and contingency. *The Monist* 62(2): 161-189. https://doi.org/10.5840/monist197962220

Evans G (1982) *The Varieties of Reference*. Oxford University Press, Oxford.

Ferguson HJ, Breheny R (2011) Eye movements reveal the time-course of anticipating behaviour based on complex, conflicting desires. *Cognition* 119(2): 179-196. https://doi.org/10.1016/j.cognition.2011.01.005

Field H (2009) *Discovering Statistics Using SPSS* (3rd Edition). SAGE, London.

Fitch GW (2004) *Saul Kripke*. Acumen, Chesham.

Frege G (1892) Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100(1), 25-50. English Translation "On sense and reference", in *Translations from the Philosophical Writings of Gottlob Frege*, Geach P, Black M (eds) pp. 56-79. Oxford: Blackwell, 1960.

García-Carpintero M (2023) Saul Kripke's groundbreaking philosophy. *Teorema: Revista Internacional de Filosofía* 42(1): 171-181.

Genone J (2012) Theories of reference and experimental philosophy. *Philosophy Compass* 7(2): 152-163. https://doi.org/10.1111/j.1747-9991.2011.00459.x

Genone J, Lombrozo T (2012) Concept possession, experimental semantics, and hybrid theories of reference. *Philosophical Psychology* 25(5): 717-742. https://doi.org/10.1080/09515089.2011.627538

Gernsbacher M, Kaschak M (2003) Language comprehension. In *Encyclopedia of Cognitive Science*, Vol 2, Nadel L (ed) pp. 723-725. London: Nature Publishing Group. https://doi.org/10.1002/0470018860.s00560

Gettier E (1963) Is justified true belief knowledge? *Analysis* 23(6): 121-123. https://doi.org/10.1093/analys/23.6.121

Goldman A, O'Connor C (2021) Social epistemology. *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Zalta EN (ed), <https://plato.stanford.edu/archives/win2021/entries/epistemology-social/>.

Grau C (2010) Love and history. *The Southern Journal of Philosophy* 48(3): 246-271. https://doi.org/10.1111/j.2041-6962.2010.00030.x

Grau C, Pury CL (2014). Attitudes towards reference and replaceability. *Review of Philosophy and Psychology* 5(2): 155-168. https://doi.org/10.1007/s13164-013-0162-3

Guirado MA (2020) The ambiguity problem in experimental semantics. *Teorema: Revista Internacional de Filosofía* 39(3): 17-34.

Häggqvist S, Wikforss Å (2015) Experimental semantics: the case of natural kind terms. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 109-138. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-006

Hansen N (2015) Experimental philosophy of language. *Oxford handbooks online*. https://doi.org/10.1093/oxfordhb/9780199935314.013.53

Haukioja J, Nyquist M, Jylkkä J (2021) Reports from Twin Earth: Both deep structure and appearance determine the reference of natural kind terms. *Mind & Language* 36(3): 377-403. https://doi.org/10.1111/mila.12278

Haukioja J, Toorman J, Baggio G, Jylkkä J (2023) Are natural kind terms ambiguous?. *Cognitive Science* 47(9): e13335. https://doi.org/10.1111/cogs.13335

Hazlett A (2010) The Myth of Factive Verbs. *Philosophy and Phenomenological Research* 80(3): 497-522. https://doi.org/10.1111/j.1933-1592.2010.00338.x

Heck RG Jr (2018) Speaker's reference, semantic reference, and intuition. *Review of Philosophy and Psychology* 9(2): 251-269. https://doi.org/10.1007/s13164-017-0362-3

Holton R (1997) Some telling examples: A reply to Tsohatzidis. *Journal of Pragmatics* 28(5): 625-628. https://doi.org/10.1016/S0378-2166(96)00081-1

Horvath J, Koch S (2021) Experimental philosophy and the method of cases. *Philosophy Compass* 16(1): e12716. https://doi.org/10.1111/phc3.12716

Howitt D, Cramer D (2014) Introduction to Statistics in Psychology (6th Edition). PEARSON, Harlow.

Ichikawa J, Maitra I, Weatherson B (2012) In defense of a Kripkean dogma. *Philosophy and Phenomenological Research* 85(1): 56-68. https://doi.org/10.1111/j.1933-1592.2010.00478.x

Ichikawa JJ, Steup M (2018) The analysis of knowledge. *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Zalta EN (ed), <https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/>.

Islam F, Baggio G (2020) Kripkeans of the world, unite!. *Journal of Semantics* 37(2): 297-309. https://doi.org/10.1093/jos/ffaa001

Jackman H (2009) Semantic intuitions, conceptual analysis, and cross-cultural variation. *Philosophical Studies* 146(2): 159-177. https://doi.org/10.1007/s11098-008-9249-6

Jeshion R (2004) Descriptive descriptive names. In *Descriptions and Beyond*, Reimer M, Bezuidenhout A (eds) pp. 591-612. New York: Oxford University Press. https://doi.org/10.1093/oso/9780199270514.003.0019

Johnson M, Deutsch M (2021) Fake gnus! (Or: there is no experimental evidence for the lazy person's approach to philosophy). *Inquiry*. https://doi.org/10.1080/0020174X.2021.1977995

Jylkkä J, Railo H, Haukioja J (2009) Psychological essentialism and semantic externalism: Evidence for externalism in lay speakers' language use. *Philosophical Psychology* 22(1): 37-60. https://doi.org/10.1080/09515080802703687

Karabanov A, Bosch P, König P (2007) Eye tracking as a tool to investigate the comprehension of referential expressions. In *Roots: Linguistics in Search of its Evidential Base*, Featherston S, Sternefeld W (eds) pp. 207-266. Berlin: DeGruyter.

Keysar B, Barr DJ, Balin JA, Brauner JS (2000) Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychological Science* 11(1): 32-38. https://doi.org/10.1111/1467-9280.00211

Knobe J (2003) Intentional action and side effects in ordinary language. *Analysis* 63(3): 190-194. https://doi.org/10.1093/analys/63.3.190

Knobe J (2022) Personal identity and dual character concepts. In *Experimental Philosophy of Identity and the Self*, Tobia K (ed) pp. 49-70. London: Bloomsbury Publishing. https://doi.org/10.5040/9781350246928.0009

Knobe J, Nichols S (2017) Experimental philosophy. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Zalta EN (ed), <https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy/>.

Koch S, Wiegmann A (2022) Folk intuitions about reference change and the causal theory of reference. *Ergo an Open Access Journal of Philosophy* 8(25): 31-57. https://doi.org/10.3998/ergo.2226

Kraut R (1986) Love de re. *Midwest Studies in Philosophy* 10(1): 413-30. https://doi.org/10.1111/j.1475-4975.1987.tb00549.x

Kripke S (1971) Identity and necessity. In *Identity and Individuation*, Munitz MK (ed) pp. 135-164. New York: New York University Press.

Kripke S (1972/1980) *Naming and Necessity*. Basil Blackwell, Oxford.

Kripke S (1977) Speaker's reference and semantic reference. *Midwest Studies in Philosophy* 2(1): 255-276. https://doi.org/10.1111/j.1475-4975.1977.tb00045.x

Kripke S (1979) A Puzzle about belief. In *Meaning and Use*, Margalit A (ed) pp. 239-283. Dordrecht: Reidel. https://doi.org/10.1007/978-1-4020-4104-4_20

Kripke S (2013) *Reference and Existence: The John Locke Lectures*. Oxford University Press, New York. https://doi.org/10.1093/acprof:oso/9780199928385.001.0001

Kroon FW (1987) Causal descriptivism. *Australasian Journal of Philosophy* 65(1): 1-17. https://doi.org/10.1080/00048408712342731

Krosnick J (1999) Survey research. *Annual Review of Psychology* 50(1): 537-567. https://doi.org/10.1146/annurev.psych.50.1.537

Kuhn T (1962/1970) *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.

Lam B (2010) Are Cantonese speakers really descriptivists? Revisiting cross-cultural semantics. *Cognition* 115(2): 320-329. https://doi.org/10.1016/j.cognition.2009.12.018

Lewis D (1994) Reduction of mind. In *A Companion to the Philosophy of Mind*, Guttenplan S (ed) pp. 412-431. Oxford: Blackwell.

Li J (2021) The origin of cross-cultural differences in referential intuitions: perspective taking in the Gödel Case. *Journal of Semantics* 38(3): 415-440. https://doi.org/10.1093/jos/ffab010

Li J (2023a) *The Referential Mechanism of Proper Names*. Routledge, Abingdon and New York. https://doi.org/10.4324/9781003047292

Li J (2023b) Cross-cultural variation and perspectivalism: Alignment of two red herrings?. *Mind & Language* 38(4): 1157-1163. https://doi.org/10.1111/mila.12470

Li J, Liu L, Chalmers E, Snedeker J (2018) What is in a name: the development of cross-cultural differences in eeferential intuitions. *Cognition* 171: 108-111. https://doi.org/10.1016/j.cognition.2017.10.022

Ludwig K (2007) The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy* 31(1): 128-159. https://doi.org/10.1111/j.1475-4975.2007.00160.x

Luk KKS, Xiao WS, Cheung H (2012) Cultural effect on perspective taking in Chinese–English bilinguals. *Cognition* 124(3): 350-5. https://doi.org/10.1016/j.cognition.2012.05.016

Machery E (2011a) Thought experiments and philosophical knowledge. *Metaphilosophy* 42(3): 191-214. https://doi.org/10.1111/j.1467-9973.2011.01700.x

Machery E (2011b) Variation in Intuitions about reference and ontological disagreements. In *A Companion to Relativism*, Hales SD (ed) pp. 118-136. New York: Blackwell. https://doi.org/10.1002/9781444392494.ch7

Machery E (2012a) Expertise and intuitions about reference. *THEORIA. An International Journal for Theory, History and Foundations of Science* 27(1): 37-54. https://doi.org/10.1387/theoria.3482

Machery E (2012b) Semantic epistemology: A brief response to Devitt. *THEORIA. An International Journal for Theory, History and Foundations of Science* 27(2): 223-227. https://doi.org/10.1387/theoria.6223

Machery E (2014) What is the significance of the demographic variation in semantic intuitions?. In *Current Controversies in Experimental Philosophy*, Machery E, O'Neill E (eds) pp. 3-16. New York and Abingdon: Routledge.

Machery E (2015) A Rylean argument against reference. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 65-83. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-004

Machery E (2021) Cross-cultural semantics at 15. In *The Routledge Handbook of Linguistic Reference*, Biggs S, Geirsson H (eds) pp. 535-550. New York and Abingdon: Routledge.

Machery E (2023) Why variation matters to philosophy. *Res Philosophica* 100(1): 1-22. https://doi.org/10.11612/resphil.2264

Machery E (2024) Experimental philosophy of language: proper names and predicates. In *The Compact Compendium of Experimental Philosophy*, Bauer AM, Kornmesser S (eds) pp. 183-210. Berlin and Boston: Walter de Gruyter GmbH. https://doi.org/10.1515/9783110716931-009

Machery E, Deutsch M, Mallon R, Nichols S, Sytsma J, Stich SP (2010) Semantic intuitions: Reply to Lam. *Cognition* 117(3): 361-366. https://doi.org/10.1016/j.cognition.2010.08.016

Machery E, Grau C, Pury CL (2020) Love and power: Grau and Pury (2014) as a case study in the challenges of X-Phi replication. *Review of Philosophy and Psychology* 11(4): 995-1011. https://doi.org/10.1007/s13164-020-00465-x

Machery E, Mallon R, Nichols S, Stich SP (2004) Semantics, cross-cultural style. *Cognition* 92(3): B1-B12. https://doi.org/10.1016/j.cognition.2003.10.003

Machery E, Mallon R, Nichols S, Stich SP (2013) If folk intuitions vary, then what?. *Philosophy and Phenomenological Research* 86(3): 618-635. https://doi.org/10.1111/j.1933-1592.2011.00555.x

Machery E, Olivola CY, Cheon H, Kurniawan IT, Mauro C, Struchiner N, Susianto H (2023) Is identity essentialism a fundamental feature of human cognition?. *Cognitive Science* 47(5): e13292. https://doi.org/10.1111/cogs.13292

Machery E, Olivola CY, de Blanc M (2009) Linguistic and metalinguistic intuitions in the philosophy of language. *Analysis* 69(4): 689-694. https://doi.org/10.1093/analys/anp095

Machery E, Stich SP (2012) The role of experiment. In *The Routledge Companion to the Philosophy of Language*, Fara DG, Russell G (eds) pp. 495-512. New York and Abingdon: Routledge.

Machery E, Stich SP, Rose D, Alai M, Angelucci A, Berniunas R, Buchtel EE, Chatterjee A, Cheon H, Cho IR, et al. (2017) The Gettier intuition from South America to Asia. *Journal of Indian Council of Philosophical Research* 34(3): 517-541. https://doi.org/10.1007/s40961-017-0113-y

Machery E, Stich SP, Rose D, Chatterjee A, Karasawa K, Struchiner N, Sirker S, Usui N, Hashimoto T (2015) Gettier across cultures. *Nous* 51(3): 645-664. https://doi.org/10.1111/nous.12110

Machery E, Sytsma J, Deutsch M (2015) Speaker's reference and cross-cultural semantics. In *On Reference*, Bianchi A (ed) pp. 62-76. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198714088.003.0004

Mallon R (2007) Arguments from reference and the worry about dependence. *Midwest Studies in Philosophy* 31(1): 160-183. https://doi.org/10.1111/j.1475-4975.2007.00155.x

Mallon R, Machery E, Nichols S, Stich SP (2009) Against arguments from reference. *Philosophy and Phenomenological Research* 79(2): 332-356. https://doi.org/10.1111/j.1933-1592.2009.00281.x

Martí G (2009) Against semantic multi-culturalism. *Analysis* 69(1): 42-48. https://doi.org/10.1093/analys/ann007

Martí G (2012) Empirical data and the theory of reference. In *Reference and Referring: Topics in Contemporary Philosophy*, Kabasenche WP, O'Rourke M, Slater MH (eds) pp. 63-82. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9581.003.0004

Martí G (2014) Reference and experimental semantics. In *Current Controversies in Experimental Philosophy*, Machery E, O'Neill E (eds) pp. 17-26. New York and Abingdon: Routledge.

Martí G (2015) General terms, hybrid theories, and ambiguity. A discussion of some experimental results. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 157-172. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-008

Martí G (2017) Names, descriptions and causal descriptions. Is the magic gone?. *Topoi* 39(2): 357-365. https://doi.org/10.1007/s11245-017-9525-1

Martí G (2020) Experimental semantics, descriptivism, and anti-descriptivism. Should we endorse referential pluralism? In *Language and Reality from a Naturalistic Perspective: Themes from Michael Devitt*, Bianchi A (ed) pp. 329-341. Cham: Springer. https://doi.org/10.1007/978-3-030-47641-0_16

Maynes J (2015). Interpreting intuition: experimental philosophy of language. *Philosophical Psychology*, 28(2), 260-78. https://doi.org/10.1080/09515089.2013.815987

McGrath M, Frank D (2020) Propositions. *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Zalta EN (ed), <https://plato.stanford.edu/archives/win2020/entries/propositions/>.

Milojević M (2023) Intuitions and theories of reference. *Theoria, Beograd* 66(3): 5-19. https://doi.org/10.2298/THEO2303005M

Nado J (2016) Introduction. In *Advances in Experimental Philosophy and Philosophical Methodology*, Nado J (ed) pp. 1-10. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474223232.ch-001

Nado J, Johnson M (2016) Intuitions and the theory of reference. In Advances in *Experimental Philosophy and Philosophical Methodology*, Nado J (ed) pp. 5-154. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474223232.ch-007

Nichols S, Pinillos Á, Mallon R (2016) Ambiguous reference. *Mind* 125(497): 145-175. https://doi.org/10.1093/mind/fzv196

Nisbett RE (2003) *The Geography of Thought: How Asians and Westerners Think Differently... and Why*. Free Press, New York.

Nisbett RE, Peng K, Choi I, Norenzayan A (2001) Culture and systems of thought: holistic vs. analytic cognition. *Psychological Review* 108(2): 291-310. https://doi.org/10.1037/0033-295x.108.2.291

Norenzayan A, Smith E, Kim B (2002) Cultural preferences for formal versus intuitive reasoning. *Cognitive Science* 26(5): 653-684. https://doi.org/10.1207/s15516709cog2605_4

Ostertag G (2013) The 'Gödel' effect. *Philosophical Studies* 166(1): 65-82. https://doi.org/10.1007/s11098-012-0021-6

Pérez Otero M (2017) Teorías de la referencia, filosofía experimental y calibración de intuiciones. *THEORIA. An International Journal for Theory, History and Foundations of Science* 32(1): 41-62. https://doi.org/10.1387/theoria.15463

Persson I (2005) *The Retreat of Reason: A Dilemma in the Philosophy of Life*. Oxford University Press, New York. https://doi.org/10.1093/0199276900.001.0001

Pickering MJ (2003) Parsing. In *Encyclopedia of Cognitive Science*, Vol. 3, Nadel L. (ed) pp. 462-465. London: Nature Publishing Group. https://doi.org/10.1002/0470018860.s00255

Pinillos A (2015) Ambiguity and referential machinery. In *Advances in Experimental Philosophy of Language*, Haukioja J (ed) pp. 139-156. London: Bloomsbury Publishing. https://doi.org/10.5040/9781474219815.ch-007

Putnam H (1975) The meaning of "meaning". In *Mind, Language and Reality: Philosophical Papers*, Vol. 2, Putnam H (ed) pp. 215-271. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511625251.014

Reimer M (2004a) Descriptively introduced names. In *Descriptions and Beyond*, Reimer M, Bezuidenhout A (eds) pp. 613-650. New York: Oxford University Press. https://doi.org/10.1093/oso/9780199270514.003.0020

Reimer M (2004b) What malapropisms mean: A reply to Donald Davidson. *Erkenntnis* 60(3): 317-334. https://doi.org/10.1023/B:ERKE.0000023383.38025.cd

Reimer M (2010) Reference. *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Zalta EN (ed) <http://plato.stanford.edu/archives/spr2010/entries/reference/>.

Rorty AO (1986) The historicity of psychological attitudes: love is not love which alters not when it alteration finds. *Midwest Studies in Philosophy* 10(1): 399-412. https://doi.org/10.1111/j.1475-4975.1987.tb00548.x

Rose D, Machery E, Stich S, Alai M, Angelucci A, Berniūnas R, Buchtel EE, Chatterjee A, Cheon H, Cho IR, et al. (2017) Nothing at stake in knowledge. *Noûs* 53(1): 224-247. https://doi.org/10.1111/nous.12211

Rowlands M (2020) Externalism about the mind. *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Zalta EN (ed), <https://plato.stanford.edu/archives/win2020/entries/content-externalism/>.

Russell B (1905) On denoting. *Mind* 14(56): 479-93. https://doi.org/10.1093/mind/fzi873

Russell B (1918/1919) The philosophy of logical atomism. *The Monist*, 28: 495-527; 29: 32-63, 190-222, 345-80. Republished as *The Philosophy of Logical Atomism*, Routledge, Abingdon. https://doi.org/10.4324/9780203864777

Russell B (1919) *Introduction to Mathematical Philosophy*. Allen and Unwin, London.

Santambrogio M (2014) Did Madagascar undergo a change in referent? In *Mind, Values, and Metaphysics: Philosophical Essays in Honor of Kevin Mulligan*, Vol. 2, Reboul A (ed) pp. 487-501. Heidelberg: Springer. https://doi.org/10.1007/978-3-319-05146-8_28

Schlesewsky M, Friederici AD (2003) Sentence processing, mechanisms. In *Encyclopedia of Cognitive Science*, Vol. 3, Nadel L. (ed) pp. 1149-1155. London: Macmillan Reference. https://doi.org/10.1002/0470018860.s00260

Searle JR (1958) Proper Names. *Mind* 67(266): 166-173. https://doi.org/10.1093/mind/lxvii.266.166

Sperber D, Wilson D. (1995). *Relevance. Communication and cognition* (2nd ed.). Blackwell, Oxford.

Sterelny K (2020), Michael Devitt, cultural evolution and the division of linguistic labour. In *Language and Reality from a Naturalistic Perspective: Themes from Michael Devitt*, Bianchi A (ed) pp. 173-189. Cham: Springer. https://doi.org/10.1007/978-3-030-47641-0_9

Strawson PF (1950) On referring. *Mind* 59(235): 320-344. https://doi.org/10.1093/mind/lix.235.320

Strawson PF (1959) *Individuals: An Essay in Descriptive Metaphysics*. Methuen, London.

Stoet G (2010) PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods* 42(4): 1096-1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet G (2017) PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology* 44(1): 24-31. https://doi.org/10.1177/0098628316677643

Sytsma J, Livengood J (2011) A new perspective concerning experiments on semantic intuitions. *Australasian Journal of Philosophy* 89(2): 315-332. https://doi.org/10.1080/00048401003639832

Sytsma J, Livengood J, Sato R, Oguchi M (2015) Reference in the land of the rising sun: A cross-cultural study on the reference of proper names. *Review of Philosophy and Psychology* 6(2): 213-230. https://doi.org/10.1007/s13164-014-0206-3

Tanenhaus MK (2003). Sentence processing. In *Encyclopedia of Cognitive Science*, Vol. 3, Nadel L (ed) pp. 1142-1148. London: Nature Publishing Group.

Tanenhaus MK, Spivey-Knowlton MJ (1996) Eye-tracking. In *Language and Cognitive Processes: A Guide to Spoken Word Recognition Paradigms*, Vol. 11, Grosjean F, Frauenfelder U (eds) pp. 583-588. Hove: Psychology Press. https://doi.org/10.1080/016909696386971

Tobia K, Newman GE, Knobe J (2020) Water is and is not H2O. *Mind & Language* 35(2): 183-208. https://doi.org/10.1111/mila.12234

Turri J (2016) Knowledge judgments in "Gettier" cases. In *A Companion to Experimental Philosophy*, Sytsma J, Buckwalter W (eds) pp. 337-348. Oxford: Blackwell. https://doi.org/10.1002/9781118661666.ch23

Ulatowski J (2018) Is there a commonsense semantic conception of truth? *Philosophia* 46(2): 487-500. https://doi.org/10.1007/s11406-017-9941-x

van Dongen N, Colombo M, Romero F, Sprenger J (2021) Intuitions about the reference of proper names: a meta-analysis. *Review of Philosophy and Psychology* 12(4): 745-774. https://doi.org/10.1007/s13164-020-00503-8

Vignolo M, Domaneschi F (2018) Referential intuitions are still problematic. *Analysis* 78(3): 472-483. https://doi.org/10.1093/analys/anx148

Vignolo M, Domaneschi F (2022) Intuitions on semantic reference. *Review of Philosophy and Psychology* 13(3): 755-778. https://doi.org/10.1007/s13164-021-00560-7

Weinberg JM (2017) What is negative experimental philosophy good for?. In *The Cambridge Companion to Philosophical Methodology*, D'Oro G, Overgaard S (eds) pp. 161-183. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316344118.010

Weinberg JM, Nichols S, Stich S (2001) Normativity and epistemic intuitions. *Philosophical Topics* 29(1/2): 429-460. https://doi.org/10.5840/philtopics2001291/217

Wikforss Å (2017) Semantic intuitions and the theory of reference. *Teorema: Revista Internacional de Filosofía* 36(3): 95-116.

Wu S, Dale JB, Gann TM, Keysar B (2013). How culture influences perspective taking: differences in correction, not integration. *Frontiers in Human Neuroscience* 7(822): 1-7. https://doi.org/10.3389/fnhum.2013.00822

Wu S, Keysar B (2007) The effect of culture on perspective taking. *Psychological Science* 18(7): 600-606. https://doi.org/10.1111/j.1467-9280.2007.01946.x

Webography

Cambridge Dictionary, "Carburetor", <https://dictionary.cambridge.org/dictionary/english/carburetor>, consulted on December 18th, 2023.

IBM, "Cramér's V", <https://www.ibm.com/docs/en/cognos-analytics/12.0.0?topic=terms-cramrs-v>, consulted on December 18th, 2023.

Wikipedia, "Jack the Ripper", <https://en.wikipedia.org/wiki/Jack_the_Ripper>, consulted on December 18th, 2023.

Youtube, "Genoveva Marti and Edouard Machery. Conversation on Reference and Experimental Philosophy", Philvideos, <https://www.youtube.com/watch?v=a_DM292-JUs>, consulted on December 18th, 2023.

Nicolò D'Agruma