

Adrian Ziółkowski¹

The Reference of Proper Names, Semantic Intuitions and Experimental Philosophy²

Abstract This paper is a contribution to the debate concerning the kind of philosophical conclusions that can (or cannot) be derived from systematic empirical studies of intuitions about the reference of proper names. The focus of the paper is the famous study by Machery et al. (2004) in which intercultural differences in semantic intuitions between American and Chinese participants were observed. Machery *et al.* used the obtained results to question the usefulness of intuitions in philosophical discussions concerning the reference of proper names.

In this paper, I present the results of my own philosophical-experimental studies aimed at analyzing methods used in research dedicated to the problem of reference rather than semantic intuitions as such. These results indicate a significant instability of responses regarding the reference of proper names and their susceptibility to the impact of philosophically insignificant factors. Based on the collected data, I argue that methods used in experimental studies concerning the reference of proper names conducted to date do not guarantee the assessment of intuitions of the desired kind.

Keywords proper names, reference, intuitions, the descriptivist theory of names, the causal-historical theory of names, experimental philosophy

¹Department of Philosophy and Sociology, University of Warsaw.

E-mail: adrian.a.ziolkowski@uw.edu.pl

²The project was financed with a grant from the National Science Centre, DEC-2012/05/HIS/03303.

Introduction

In recent years, philosophers have begun to advance arguments of a new kind. These arguments make significant references to empirical data gathered systematically for this purpose. Philosophers using such methods – experimental philosophers, as they call themselves – claim that many theses put forth in “classical” philosophical debates are empirical in nature and should be tested accordingly. At issue here is, first and foremost, the sometimes explicitly formulated expectation of uniformity of intuitions about popular philosophical thought experiments. As part of their research, experimental philosophers attempt to determine: (1) if the intuitions of nonphilosophers track those most often expressed in the literature (e.g. intuitions regarding the ascription of knowledge in Gettier scenarios); (2) what factors might contribute to the formation of such intuitions, and (3) if these intuitions can be systematically distorted.

Experimental philosophy is a young field: its history as an organized research enterprise in analytic philosophy goes back to the beginning of the 21st century³. Since then, experimental philosophers have sought to apply their methods in almost every philosophical discussion significantly shaped by thought experiments. This paper focuses on experimental methods used in the study of the reference of proper names, although this is not the only problem in the philosophy of language addressed by experimental philosophers to date – another example are intuitions about the contextual dependency of natural language expressions. Experimental philosophers have also participated in discussions pertaining to epistemology (of particular interest here has been the problem of knowledge ascription), moral philosophy (e.g. in relation to the famous trolley problem), action theory (in regard to questions such as intentionality or free will), philosophy of mind (focusing on the problem of consciousness, especially its types), and ontology (here, the intuitions of nonphilosophers about causation have been studied, among other issues). According to experimental philosophers, their methods can be applied to any problem addressable using thought experiments, as long as these can be presented in the form of scenarios comprehensible to nonphilosophers (an overview of issues addressed during the early stages of the development of experimental philosophy can be found in (Alexander, 2012)).

³In 2008, the leading figures of this emerging field published *An Experimental Philosophy Manifesto*, defining the field’s main goals and research approach (Knobe & Nichols, 2008).

Experimental philosophy is not a homogenous field, a range of different conclusions are derived based on data obtained during research. Experimental philosophers often claim that data concerning the intuitions of nonphilosophers support or refute certain philosophical theories – for example, the fact that respondents are disinclined to ascribe knowledge to the protagonists of Gettier scenarios has been seen as an argument against the classical definition of knowledge. That said, researchers are sometimes more interested in undermining the overall usefulness of intuitions (or thought experiments) in a given field by demonstrating that responses elicited by thought experiments are systematically distorted – if it turns out that the intuitions being tested depend on some irrelevant factor, such as the respondent’s ethnic background, they ought not to be trusted. It might be worth noting that these two ways of utilizing empirical data in experimental philosophy are almost antithetical. Therefore, philosophical-experimental research of the first kind is sometimes referred to as positive experimental philosophy, and research of the second kind, as negative experimental philosophy (Nadelhoffer, Nahmias 2007).

The legitimacy of the mode of philosophizing proposed by experimental philosophers is the subject of lively debate. Its beginning can be traced to the publication by Edouard Machery, Ron Mallon, Shaun Nichols, and Stephen Stich⁴ (2004) of an article titled *Semantics, cross-cultural style*, the focus of this paper⁵. The article, written in a provocative style, garnered much scholarly interest as its authors used novel methods to formulate, in the spirit of negative experimental philosophy, strong conclusions that cast doubt on other, more “classical” philosophical approaches. The main research goal of MMNS was to gather empirical data which could be used to undermine the usefulness of semantic intuitions in debates concerning the reference of proper names. They searched for factors that seem philosophically insignificant but nonetheless impact semantic intuitions – their data, indicating cross-cultural differentiation in semantic intuitions about the reference of proper names, suggested that cultural background is one such factor.

⁴Hereafter referred to as MMNS.

⁵This is, to be precise, not the first publication considered to belong to the field of experimental philosophy. One earlier example is the report on the results of research on epistemological intuitions about Gettier scenarios and other related thought experiments by Weinberg, Nichols, and Stich (2001). The obtained data indicated that intuitions of this kind depend on the ethnic background and socioeconomic status of the respondents. However, the reliability of these results is doubtful – the experiment was conducted on very small respondent groups, and the more recent attempts to replicate the study turned out unsuccessful (Kim & Yuan, 2015; Seyedsayamdost, 2015).

In this paper, I present the results of my own studies. They cast doubt on the efficacy of research tools used to date to reveal the semantic intuitions of nonphilosophers, thus undermining the controversial thesis put forth by MMNS. According to MMNS, cultural differences in responses to the presented scenarios they observed testify to differences in semantic intuitions. The results of my experiments, methodologically modelled on the study conducted by MMNS, can be interpreted, as I am going to argue, to indicate that experimental methods used in this kind of research do not guarantee the acquisition of data reflecting intuitions of the desired kind. And since cultural differences noted by MMNS need not mean differences in semantic intuitions, the impact of their attack on the appeal to intuitions in debates concerning the reference of proper names is diminished.

In the first part of the paper, I briefly address controversies surrounding the notion of intuitions in philosophy. In the second part, I discuss the results of the study by Machery et al. indicating, in their opinion, the uselessness of semantic intuitions in philosophical debates. In the third part, I briefly recount the most important objections to their methods and interpretations to be found in the literature to date. In the fourth and fifth parts, I outline my own critique backed by the results of my experiments.

1. Intuitions in philosophy

The problem of intuitions has been the subject of considerable controversy in contemporary philosophical literature. It is often claimed that intuitions – particularly conceptual intuitions – are the factor responsible for our responses to philosophical thought experiments. On the other hand, no small number of philosophers have questioned the role of intuitions in philosophical debates. Timothy Williamson (2007), for example, claims that thought experiments are in fact modal arguments and that intuitions are irrelevant to the soundness of these arguments, thus playing no role in philosophical debates. Herman Cappelen (2013) has formulated an even more categorical thesis; according to him, the use of the term “intuition” in philosophy is so vague that it should be concluded that intuitions do not exist.

The crucial *locus* of conflict, regardless of the stance on the role of intuitions in philosophical considerations one might be inclined to adopt, is their nature – the only point philosophers agree on is that intuitions are mental states of a special kind. At least five individuation criteria for intuitions have been proposed in the scholarship so far, appealing to: (1) their phenomenal character; (2) their content; (3) their epistemic status; (4) their

origin; and (5) their functional role. The first approach suggests that the distinguishing characteristics of intuitions are their phenomenal properties – it is claimed, for instance, that intuitions are “seemings” accompanied by the impression that their content is necessarily true (e.g. Bealer, 1998). The content criterion may point to the abstract nature of the contents of intuitions (intuitions are supposed to not concern contingent objects) or to their modal nature (they are supposed to adjudicate questions of possibility or necessity). According to the third approach, the characteristic feature of intuitions is that they possess an *a priori* justification. The origin criterion refers to the cognitive competency considered to be the source of intuitions – usually linguistic competency and its related capacity to comprehend certain concepts (e.g. Ludwig, 2007 and 2010). The last approach to the problem of intuitions mentioned above is to determine some specific functional property shared by intuitions such as the fact that they are spontaneous reactions to philosophical thought experiments.

As can easily be seen, some of the aforementioned criteria are relatively strict and some relatively liberal – on some (e.g. the functional criterion), numerous mental states are going to be classed as intuitions; on others (e.g. the origin criterion), the set of mental states identifiable as intuitions is going to be considerably smaller. The solution of this conflict is beyond the purview of this paper. For the sake of the current discussion, I propose to assume a liberal, functional characterization of intuitions as spontaneous reactions to thought experiments – such characterization is presumed by the majority of experimental philosophers in their research (Weinberg & Alexander, 2014). This leaves open the question of whether a given spontaneous reaction to a scenario presented in an experimental study is an intuition of the desired kind, that is, if it is relevant to the philosophical issue raised in the study. The fundamental question I address in this paper is this: do methods hitherto employed in experiments concerning the problem of the reference of proper names provide data reflecting the sought-after intuitions, that is, intuitions expressing support for the particular theory of reference? As I have noted earlier, I am going to use the results of my research to argue for a negative answer to this question.

2. The study by Machery *et al.*

2.1. The area of dispute: two competing theories of reference

Let us discuss the results of the research to date. The departure point of the study conducted by MMNS (2004) was a debate between two historically

strongest traditions of thinking about the reference of proper names: the descriptivist and the causal-historical tradition. Neither of these traditions can be considered perfectly homogeneous; rather, they are groups of theories that share a common core but differ in detail. For simplicity's sake, I am going to speak of the descriptivist and the causal-historical theory of reference. Below, I present an outline of the key assumptions made in these approaches according to their characterization introduced into the literature by Machery et al. It is important to note that this outline paints a general, simplified and imprecise picture of two conceptions of reference neither of which deserves to be called a theory. Views on the problem of reference held by particular philosophers of language tend to be much more precise. However, even though the following reconstruction can be considered inadequate, it is important to present it in this form since it is this articulation that has shaped the approach to the problem of the reference of proper names dominant in experimental philosophy to date. Experimental philosophers seek to empirically distinguish between intuitions providing support precisely for such general, imprecisely characterized conceptions⁶.

According to the descriptivist theory of reference, as it is reconstructed by MMNS, proper names are strictly tied to descriptions fulfilling two requirements: (i) the object referred to by a given name satisfies the description associated with that name; and (ii) this object is the only object in the *universum* that satisfies this description. Names “pick out” their reference from extralinguistic reality by means of such descriptions. MMNS go so far as to suggest that, according to the descriptivist, proper names simply are hidden descriptions, an opinion many proponents of descriptivism are likely to reject (e.g. Searle, 1985). The conception analyzed here is closest to the classical (and slightly archaic) standpoint of Frege (1977). Frege did tend to identify names with descriptions; he also permitted the possibility that the same name can be tied to different descriptions by different language users (that is, he permitted the instability of meaning for names) as long as all relevant descriptions unequivocally identified the same object. One language user, to give an example, could tie the name “Lech Wałęsa” to

⁶This fact can serve as the basis for another argumentation strategy against the study by MMNS: if their research does not concern intuitions relating to actual theories of reference, it can be considered philosophically irrelevant. In this text, I pursue a different strategy and claim that the methods proposed by MMNS do not guarantee empirical differentiation even between such generally characterized conceptions of the reference of proper names.

the description “the first head of Solidarity,” and another to the description “the first Polish laureate of the Nobel Peace Prize.”

The proponents of the causal-historical theory of reference, as it is reconstructed by MMNS, on the other hand, reject the notion that proper names should be perceived as mediated by other lexical units and hold them to be relatively independent carriers of the relation of reference. According to this theory, the relation of reference obtaining between a name and its designate is based on the history of the continuous use of that name. This history, in the case of each name, begins with the act of initial baptism whereby an individual language user (or a group of such users) introduces a convention for the use of the name in reference to some object. The name can be introduced into the language through ostension (the utterance “May this object bear the name N” accompanied by pointing to the “baptized” object) or by means of a determinate description. However, the name is autonomous in relation to the initial act of ostension or the description used in the course of the name’s introduction into the language. The relation of reference between a proper name and its designate obtains thanks to the causal chain linking the current uses of the name to the act of initial baptism. It should be noted that this characterization of the causal-historical theory of reference is also a considerable simplification. It is hard to think of the outlined standpoint as a philosophical theory in the strict sense of the term; rather, it is a relatively general idea in need of precise articulation if it is to become a theory⁷.

The goal of the study conducted by MMNS was to reveal preferences shared by members of different cultural groups in regard to the competing conceptions of the reference of proper names characterized above. It was carried out in reference to the famous thought experiments intended to counter the descriptivist theory of reference presented by Saul Kripke in the lecture series published as *Naming and Necessity* (1972).

In the first of these counterexamples, Kripke analyzes a situation in which the users of a certain name tie it to a description that is not satisfied by the object singled out as its designate by the history of linguistic practices related to the name in question. Kripke proposes the following hypothetical scenario. People generally associate the description “the author of the proof of the incompleteness theorem” with the name “Kurt Gödel.” However, contrary to common belief, this proof is in fact due to a little known German mathematician named Schmidt. Schmidt died in unexplained circumstances

⁷For more precise and more adequate descriptions of the different philosophical theories of the reference of proper names, see Lycan (1999) and Muszyński (2000).

in Vienna and his friend, Gödel, got hold of the manuscript of the proof and published it under his own name. This raises the question: who does the user of the name “Kurt Gödel” associating this name with the description “the author of the proof of the incompleteness theorem” refer to here? To Schmidt or to the person who published the proof under their own name? The first option is supposed to obey the spirit of descriptivism; the second, the spirit of the causal-historical theory of reference. There is moderate consensus among philosophers that our semantic intuitions are bound to push us in the direction of the second option, undermining the descriptivist theory and providing support for the causal-historical conception of reference.

The second Kripkean counterexample concerns situations in which the description associated with a given name is not satisfied by any object in the universe of discourse. Kripke considers the example of the biblical prophet Jonah who was swallowed by a giant fish (or whale) for three days and three nights according to legend. Let us assume, after Kripke, that the story of the swallowing is an untrue tale, although the prophet whose life served as the basis for it really existed. This raises a question analogous to the one posed in the context of the Gödel case described above: who does the language user tying the name “Jonah” to the description “the prophet swallowed by a giant fish for three days and three nights” refer to? The causal-historical theory of reference permits the possibility of referring to the actual prophet. The same is not the case for descriptivism – since the key to the relation of reference in descriptivism is the description associated with the name, and the description under consideration is not satisfied by any object, the relation of reference simply does not obtain. In this case too, the majority of philosophers agree that our intuitions tend to support the causal-historical rather than the descriptivist theory of reference.

2.2. The experimental procedure and results of the study by Machery *et al.*

MMNS modelled the scenarios used in their study on the Kripkean thought experiments outlined above, dividing them into two types: Gödel cases and Jonah cases. Their interest focused on the so-called uniformity conjecture present in Kripke’s argumentation, in their opinion. According to this conjecture, (a) there is going to be a far-reaching agreement among ordinary language users concerning the right responses to Gödel and Jonah cases, and (b) this agreement is going to favour the causal-historical theory. Therefore, in particular, no systematic differences in intuitions among groups

defined by philosophically insignificant factors, such as cultural background, should occur.

In order to find empirical evidence against the uniformity conjecture, MMNS decided to test semantic intuitions elicited by Gödel and Jonah cases in two distinct cultural groups expected to exhibit differences of the relevant sort given some more general psychological differences observed for them in previous research. Based on the results of cross-cultural studies conducted by Richard Nisbett (2003), in which the Western and East Asian traditions had been compared, MMNS formulated the expectation that there should occur systematic differences between these groups in regard to preferences related to theories of reference. The rationale was that, as Nisbett's research implied, members of East Asian culture are much less inclined to use causal categories to formulate descriptions of reality than are persons raised in the Western cultural sphere. Since the causal-historical theory is based on an explanation in causal terms, the researchers expected that intuitions supporting the descriptivist theory would be significantly more frequent among Asians than among members of Western culture – the cultural background of the majority of contemporary analytic philosophers so readily accepting Kripke's counterexamples.

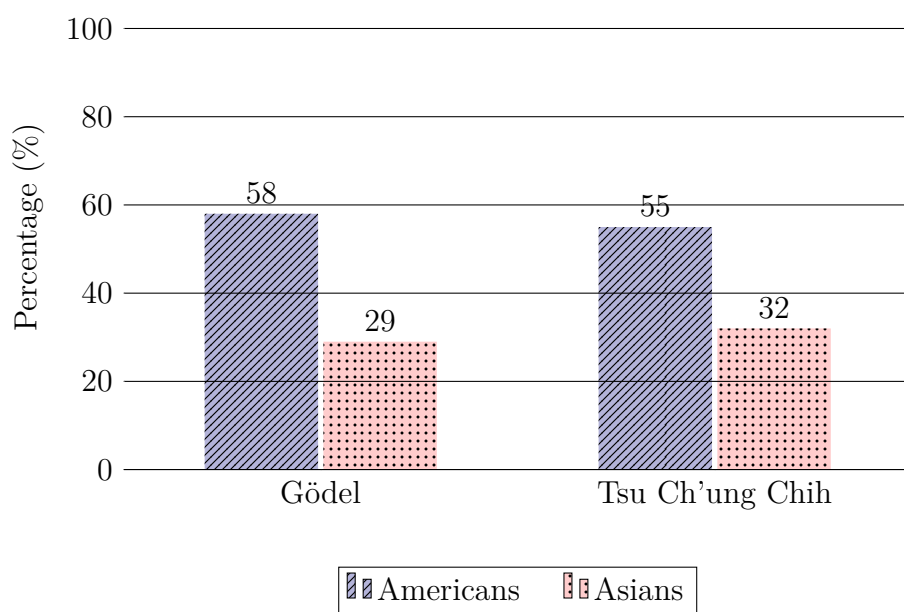
In order to verify this hypothesis, MMNS carried out an experiment involving members of Western culture (students of Rutgers University in the USA) and persons raised in the Asian cultural sphere (students of the University of Hong Kong)⁸. The groups comprised 40 persons each. The participants were presented with two scenarios of each type mentioned above: two Gödel cases and two Jonah cases. The scenarios presented to the two groups were almost identical, the main difference being that in one story of each type the protagonists bore names characteristic of Western culture (e.g. "Gödel"), and in the other, names typical of Asian culture (e.g. "Tsu Ch'ung Chih"). Importantly, the language of the experiment for both groups was English (students from Hong Kong were supposed to be fluent users of English).

The respondents revealed their preferences by choosing one of two response options. For Gödel cases, the descriptivist option was "[the protagonist] is talking about a person who actually satisfies [the description associated with N]," and the option consistent with Kripkean intuitions was "[the protagonist] is talking about a person commonly thought to satisfy [the description associated with N]." For Jonah cases, a respondent could express

⁸These were not seasoned philosophy students reasonably expected to be familiar with the problem discussed here.

their support for descriptivism by choosing the option “[the protagonist] is talking about a fictional person who does not really exist,” and their support for the causal-historical conception by choosing the other option, stating that reference to an actual person who had inspired the false myths and legends obtains. Since the respondents had no opportunity to introduce an original response to the question posed, or to pick a third option such as “I do not know” or “none of the above is correct,” one can speak of a forced choice here.

Whereas no statistically significant differences between persons raised in Western and Asian cultures were observed for Jonah cases (in both groups intuitions supporting the causal-historical theory predominated slightly), a clear difference between the groups occurred for Gödel cases: while responses supporting the descriptivist theory predominated among the Chinese, the Westerners were slightly more inclined to support the causal-historical theory. For details, see Graph 1.



Graph 1. Percentage of responses supporting the causal-historical theory for Gödel cases in the original study by Machery *et al.* (after: Machery, 2012).

It might be worth noting that although responses identified as supporting the causal-historical theory predominated among the Americans, almost half

of them reacted to Gödel cases by choosing a response considered to support the descriptivist conception. It could thus be concluded that the results for the members of Western culture alone undermine the uniformity conjecture since a considerable disagreement occurred in this group as to the correct response to Kripke's thought experiments.

The conclusion that the authors derived from the study is quite revolutionary as far as philosophical considerations regarding semantic issues are concerned. As has been mentioned earlier, MMNS claim that the data they obtained pose a serious challenge to the assumption, prevalent in philosophy, that the intuitions of philosophers about the reference of proper names (on which there is moderate consensus in philosophical circles) are universal. But this is not all. Citing the inclination on the part of Westerners, supposedly observed in their study, to support the causal-historical theory, Machery *et al.* suggest that a similar inclination among analytic philosophers, the majority of whom are Westerners, might be an expression of cultural conditioning or academic indoctrination adjusted to the demands of Western culture. And as they argue further, there is no conclusive argument for the idea that the semantic intuitions of philosophers from the West should be more accurate than, for example, those of Asians not trained in philosophy. In light of this, they propose a revision of the role assigned to semantic intuitions in contemporary philosophical discussions since the mode of doing philosophical semantics to date "smacks of narcissism in the extreme," in their opinion (Machery et al. 2004, p. B9).

3. Existing critique of the study by Machery *et al.*

The conclusions formulated by MMNS and their study have been subject to intense critique. Due to limitations of space, I cannot discuss all questions raised. In what follows, I provide a short description of the most important objections lodged by critics so far.

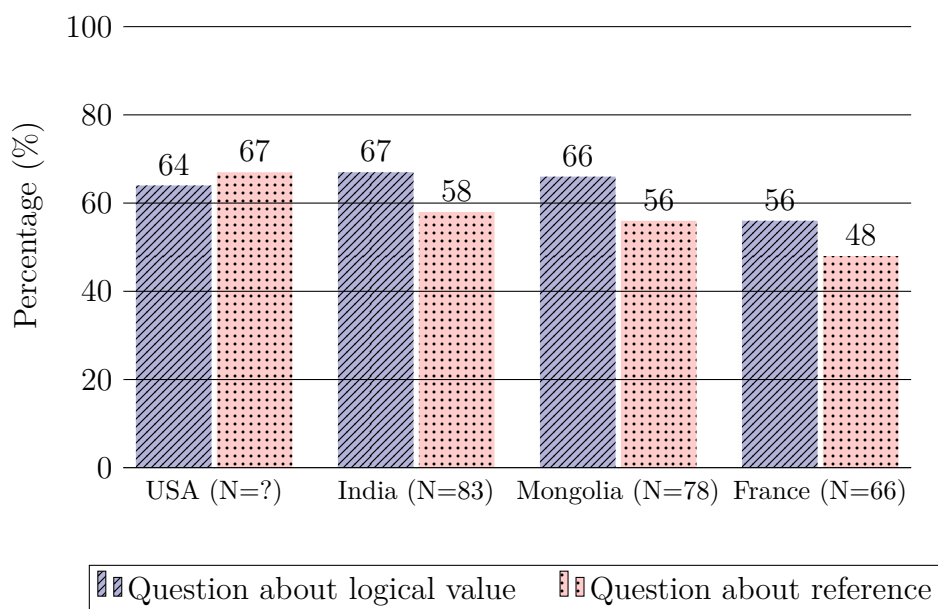
The study by MMNS, like all research in experimental philosophy, can be criticized from the position that philosophers are experts in regard to intuitions. Michael Devitt (2011) and Kirk Ludwig (2007 and 2010), for instance, hold that just like ordinary intuitions about issues in physics count less than those of experienced physicists, so too, ordinary intuitions about theories of reference count less than those of trained semanticians. Devitt (2011) also points to the excessive emphasis that MMNS put on the thought experiments they used – Gödel cases and Jonah cases – particularly in terms

of their role in Kripke's argumentation against the descriptivist theory of reference. As Devitt rightfully points out, even if MMNS's argumentation suffices to undermine the efficacy of these thought experiments, other elements of Kripke's refutation of descriptivism still retain their force and are enough to cast doubt on the criticized theory.

Since my paper focuses on methodological issues, I find such general objections much less interesting than more specific critical remarks – ones that do not undermine the legitimacy of MMNS's entire enterprise but point to elements in their experimental procedure that may have resulted in the acquisition of data related to a phenomenon altogether different from the intended one. Arguing in this spirit, Genoveva Martí (2009) doubts if the question posed to the respondents (“when the protagonist uses the name N, they are talking about. . .”) – formulated such that the key name is mentioned rather than used – actually elicits the intended semantic intuitions. Intuitions of the desired kind should be about the mechanism of reference but, according to Martí, the question might have encouraged the respondents to instead focus on the theory that best describes this mechanism. As Martí argues further, insofar as competent language users not trained in philosophy can be seen as experts in the first matter, one ought not to expect their intuitions about the correct theory of reference to be authoritative.

In response to this objection, Machery *et al.* (2009) carried out an additional study of Gödel cases to determine if significant differences in response distribution would occur between the original question and its alternative formulation asking about the logical value of a sentence where the key name is used (e.g., “Gödel is the author of the incompleteness theorem”). If a participant's response was positive, it was treated as an expression of a semantic intuition supporting the descriptivist theory; if it was negative, it was seen as a token of support for the causal-historical theory. The participants responding to the two formulations of the question belonged to separate groups. Machery *et al.* (2009) asked members of four cultural groups: persons from India, Mongolia, France, and the USA.

Contrary to Martí's suspicion (Martí, 2009), the experimenters did not observe statistically significant differences in response distributions for the alternative formulations of the question for any of the tested groups (for results, see Graph 2). According to Machery *et al.*, this is enough to dismiss the objection that these formulations elicit intuitions of different kinds. Although I consider this conclusion to be premature, I am not going to explore this issue any further here. Something else should be noted: although the results of the second study by Machery *et al.* (2009) confirm the existence



Graph 2. Percentage of responses consistent with the causal-historical theory in the study by Machery *et al.* (2009).

of cross-cultural differences in reactions to Gödel cases, they also undermine the hypothesis which was supposed to explain them. As it turns out, the French express intuitions supporting the causal-historical theory significantly less often than not only Americans but also persons from Mongolia. At the same time, no difference for one of the formulations (the assessment of the logical value of the sentence in which the key name is used) between the Americans and the Mongols was observed. Therefore, if there are in fact cross-cultural differences in intuitions about the reference of proper names, they are not systematic East West divergences – some other phenomenon must be responsible for their occurrence than the one suggested by MMNS (2004) in their original study.

Another objection, one concerning the language of the original study, has been raised by Barry Lam (2010). The objection is that both the Americans and the respondents from Hong Kong assessed scenarios presented in English, the native language of the first group only. Lam thinks that, given this, differences in responses to Gödel cases chosen by the Americans and the Chinese need not testify to differences in intuitions about the reference of proper names between the two groups. In his opinion, an equally plausible

hypothesis is that the observed differences are due to varying degrees of linguistic competency.

In order to back his suspicion, Lam (2010) carried out an experiment in which he presented the users of English and Cantonese with appropriate native language translations of a scenario modelled on Gödel cases. The respondents learnt a story according to which a certain group does not know anything about Shakespeare other than the fact that he is the author of *Romeo and Juliet*. According to the scenario, the truth is that the play was not written by Shakespeare but by an unknown German writer by the name Spencer. The data obtained by Lam differed significantly from those acquired by MMNS – responses supporting the causal-historical theory predominated in both cultural groups.

This result provoked Machery et al. (2010) to attempt to replicate their original study in two languages. This time, they compared reactions of Americans to the original Gödel scenario (featuring the name “Gödel”) to reactions of Chinese participants to a translation of this scenario. Contrary to the result obtained by Lam (2010), the cross-cultural differences observed in their original study were replicated: whereas responses supporting the causal-historical theory predominated among the Americans (62.2%), the Chinese were more inclined toward descriptivism (61%). The difference between data obtained by Lam on the one hand, and those acquired in the original experiment and the later study by Machery *et al.* (2010), on the other, might be the effect of using different scenarios or slightly different formulations of the key question regarding the name’s reference. Options available to the respondents in the original study by MMNS were descriptions; those given in Lam’s experiment were proper names – the respondent, when asked about the person referred to by the protagonist of the story using the name “Shakespeare,” could choose between the option “Shakespeare” (in support of Kripke) and “Spencer” (in support of descriptivism). As has been rightfully noted by Beebe and Undercoffer (2016), the latter design does not allow one to successfully adjudicate between the two competing theories – for the descriptivist both responses are correct since, according to descriptivism, in the presented situation the names “Shakespeare” and “Spencer” co-refer. They both designate the actual author of *Romeo and Juliet*. The results of Lam’s study should thus be approached with considerable caution.

An extremely interesting objection has been raised by Justin Sytsma and Jonathan Livengood (2011) who suggest that there is a dangerous ambiguity in the original study by MMNS. According to them, the formulation of the question regarding reference in this study did not sufficiently determine the

cognitive perspective that the respondent ought to assume while assessing the situation presented in the scenario. In particular, it is not clear if the desired perspective is one of the protagonist of the story (the user of the name ignorant of many facts, including the existence of Schmidt and his authorship of the incompleteness theorem) or one of the omniscient narrator (assuming access to information not possessed by the protagonist).

In order to test their hypothesis, Systma and Livengod (2011) decided to compare respondent reactions to three different formulations of the question regarding the reference of the name. The researchers used the original Gödel scenario (featuring the name “Gödel”) from the study by MMNS. The only change concerned the formulation of the question – the one meant to encourage the respondents to assume the perspective of the protagonist was “when John uses the name ‘Gödel,’ John thinks he is talking about. . . .”; on the other hand, to encourage the assumption of the perspective of the narrator, they used “when John uses the name ‘Gödel,’ he is in fact talking about. . . .” The participants were divided into three groups: one answered the question in its original formulation, and the other two, questions unambiguously indicating the protagonist and the narrator perspective respectively.

Systma and Livengod (2011) observed significant differences in reactions between formulations encouraging respondents to assume the perspective of the protagonist and the narrator – in the first case the majority (78%) concluded that the protagonist thinks he is referring to the person who actually proved the incompleteness theorem; in the second case the majority (57.4%) stated that the protagonist in fact referred to the person who got hold of the manuscript and published it under their own name. Moreover, there also occurred a significant divergence in response distribution between the original version of the scenario (where less than 40% of the respondents chose the causal-historical response) and the other two versions. This means that even a slight change in the content of the question (the addition of “in fact”) can translate into a significant change in response distribution. Based on this result, Systma and Livengood conclude that the method of measuring respondent opinion regarding the reference of proper names used in research to date does not provide decisive data as far as adjudicating between the competing theories of reference is concerned. It is simply not certain that all participants in the original experiment by MMNS responded to the same problem.

A possible ambiguity of the key question posed in the experiment by MMNS has also been pointed out by Kirk Ludwig (2007) and Max Deutsch (2009). In their opinion, the formulation of this question – the question about

the person the protagonist is talking about – does not distinguish between two significantly different senses of the term “reference”: speaker’s reference and semantic reference. The first concerns the person the user of a given name intends to refer to; the second concerns the actual reference of the name in the same use. The conflict between descriptivism and the causal-historical theory regards semantic reference, not speaker’s reference. According to Ludwig and Deutsch, there is a risk that some of the respondents may have understood the question asked by the experimenters differently, and that while some expressed their intuitions in regard to semantic reference, others addressed the problem of speaker’s reference. This could pose a serious challenge to MMNS since, granted the objection, interpreting all or at least some (which?) of the responses as supporting either of the competing theories might be invalid.

The discussion between Machery and Deutsch ultimately resulted in a cooperative project – they decided to join forces to experimentally test the aforementioned objection. The experiment by Machery, Deutsch and Systema (2015) used the Gödel scenario (featuring the name “Gödel”) from earlier research, but featuring a clear formulation of the question regarding reference: “when the protagonist of the story uses the name ‘Gödel,’ regardless of his intention, he is in fact talking about. . .”. Just like in the original study by MMNS, the scenario was assessed by American and Chinese participants. It turned out that the results for the clear formulation did not significantly diverge from the original study. 59.9% of the Americans and only 38.8% of the Chinese chose the response linked to the causal-historical theory. Differences between the groups once again turned out to be statistically significant, apparently dismissing the objection raised by Ludwig and Deutsch.

As can be seen from the discussion surrounding the results of the numerous studies based on methods similar to that initially employed by MMNS, there is a considerable instability in response distribution even if the material is largely similar or identical to that used in the original study. In the most recent attempt to replicate the experiment by MMNS carried out by Beebe and Undercoffer (2016) and involving a sufficiently large respondent group, effects similar to those observed by MMNS occurred: the Chinese were less inclined than the Americans to support the causal-historical conception for Gödel cases, and no cross-cultural differences occurred for Jonah cases. That said, the distributions were different than in the original study. For Gödel cases, the differences were small but significant (53% of the Americans and 43% of the Chinese chose the causal-historical option); for intuitions elicited by Jonah cases, the majority of which (ca. 2/3) in the study by

MMNS coincided with Kripke's intuitions, clearly supported descriptivism in the replication study (also ca. 2/3).

The discussion surrounding the controversial study by MMNS and their conclusions briefly outlined above does not yield an unequivocal picture. The remainder of this text is devoted to a detailed description of my own studies concerning intuitions elicited by scenarios modelled on Kripke's counterexamples. I am going to use the results obtained during these studies to argue in favour of the positions of the critics. My argumentation pertains to methodological issues – I am going to claim that the method used by MMNS does not guarantee that the assessment of participant responses reflects their semantic intuitions.

4. The author's research

4.1. Basic premises and main goals

So far, the discussion of the results of philosophical experiments concerning the reference of proper names has focused predominantly on Gödel cases, for which an interesting disproportion has been noted, and much less on Jonah cases. However, Jonah cases could provide data crucial for a deeper understanding of how ordinary intuitions about reference are shaped. According to Devitt (2011), Jonah cases can tell us more about the ordinary notion of reference because they are closer to problems typically encountered by everyday language users⁹.

My studies focused on Jonah cases. One reason for that was to fill the gap in the existing literature. The second and more important reason was my suspicion that Jonah cases had been adapted for philosophical-experimental research in a methodologically flawed way. Namely, one of the options that the respondents could choose from in the original study by MMNS (2004) – the option interpreted as expressing support for the descriptivist theory of proper names – did not constitute a response the proponent of this theory should prefer in a Jonah context. The option identified as descriptivist was: “[the protagonist] is talking about a fictional person who does not really exist.” This suggests that the name in question refers to some fictional object. It is difficult to tell exactly how the expression “fictional object” should be understood here but, regardless of its interpretation, there is no doubt that, as far as Jonah cases are concerned, the classical descriptivist theory

⁹It should be noted that, in general, Devitt has considerable reservations concerning the possibility of extracting accurate and competent semantic intuitions out of nonphilosophers using techniques proposed by experimental philosophers.

of proper names does not entail reference to a particular object. Jonah cases are situations in which a name is tied to a description that does not unequivocally designate any object in the world. The descriptivist response in a Jonah situation should thus indicate reference failure. The first person to point this out was Henry Jackman (2009). In light of this, it is doubtful that choosing the aforementioned option can be considered an expression of support for descriptivism.

The main goal of my experiments was to determine if formulating a more adequate option expressing support for descriptivism would spur different reactions from those elicited by the original response option. Inspired by suggestions made by Jackman (2009) and Deutsch (2009), I also decided to check if different Jonah scenarios constructed based on the same strategy as the story presented by MMNS would result in different responses. I thus set out to determine, first, if differences in response distribution for a given scenario would occur depending on the formulation of one of the response options, and secondly, if response distributions for different Jonah cases would turn out the same. The latter effect is to be expected if the respondents express support for one consistent theory of reference. The second venture was purely exploratory but it ended up providing data which I found the most interesting.

4.2. Design and procedure

Three Jonah type scenarios were prepared for the purpose of the study. They were similar in structure to the stories presented in the original study by MMNS. Each scenario described a language user belonging to a larger linguistic community tying a given proper name to a description. The titles of these scenarios – *Mapemba*, *Homer*, and *Einstein* – stem from the proper names used in each scenario. All three stories contain the information that, contrary to the opinion widely shared by the appropriate linguistic community, a single person satisfying the description associated with the key name never existed. This is described in the scenarios as resulting from a “mythologization” of an actual historical person (*Mapemba*, *Homer*) or a simple mistake (*Einstein*). Importantly, although in no case is the description true of some one individual and that individual only, the descriptions featuring in *Homer* and *Einstein* could be treated as nonempty general names – they could be truthfully predicated of every member of a group of authors/inventors indicated in the scenarios.

Each participant in the experiment learnt about all three scenarios presented in random order. Each respondent was randomly assigned to one

of three versions of the scenarios – *Fictional Person*, *No One in Particular*, and *No Reference* – corresponding to different formulations of the response option aspiring to reflect intuitions supporting the descriptivist theory. In the *Fictional Person* version, the formulation from the original study by MMNS was used: “[the protagonist] is talking about a fictional person who does not really exist”¹⁰; the formulation in the *No One in Particular* version was: “[the protagonist] is not talking about anyone in particular”; and the formulation in the *No Reference* version was “[the protagonist] does not refer to anyone.” The last formulation is assumed to come closest to the response that the proponent of the descriptivist theory should issue in each situation. Significantly, the response assumed to express support for the causal-historical conception was the same across all three versions of each scenario. Hence, the only manipulation consisted in changing one response option available to the respondent. The expected divergence in response distribution for the particular versions (formulations of the second response option) was thus assumed to provide an argument in favour of the thesis that the three alternative formulations of the descriptivist option¹¹ do not in fact express the same ordinary intuitions about the reference of proper names.

All participants in the study were native speakers of English. The scenarios were presented in English in the form recounted below.

4.3. The material: three Jonah cases

Below I present the content of the three scenarios prepared for the purpose of the study. Differences between their particular versions have

¹⁰As has been rightfully noted by an anonymous referee, this formulation is troublesome for an additional reason – it is a pleonasm, that is, the same thought is stated twice. The expression can be seen as unfortunate and thus should not be used since such elements lead to interpretive problems. In light of Grice’s theory of conversational maxims (e.g. 1975), tautological utterances usually constitute breaches of the maxim of quantity (they convey superfluous information) and thus suggest an occurrence of a conversational implicature (indirect communication). MMNS certainly did not intend to use this formulation to encourage the respondents to search for conversational implicatures. Unfortunately, since the goal of my study is to test respondent reactions to scenarios designed based on the same method as the one used in the experiment conducted by MMNS, I could not avoid “inheriting” this flaw from their original study.

¹¹I use the expression “descriptivist option” as shorthand – of course, if the respondents’ preferences in the case of the three supposedly descriptivist responses are not identical, one cannot say that they all support descriptivism. Strictly speaking, at least two of them are not descriptivist responses but responses initially intended to express descriptivist intuitions, although they do not actually do so.

been indicated. These differences concern questions only; the content of each scenario is the same across all versions.

MAPEMBA

Thenga lives in a small African town called Kwende. Like most of Kwende inhabitants, he believes that Kwende was founded by a shaman called Mapemba. Moreover, Thenga believes that people who lived in the area before Mapemba's reign had been affected by recurring plagues and catastrophes. In fact, Mapemba is believed to have stopped those plagues with his magical powers and thus to have given the people of Kwende a peaceful life.

The truth is different, however. There never was any shaman who fought the plagues with magical powers. The origins of Kwende are connected with the activity of an inventive tribal leader called Ndembo, who had an idea how to use a nearby river to irrigate the cultivations and increase their efficiency. This solution improved the living standards of Kwende's people so much that it gave rise to a legend. The story of the inventive leader was passed from generation to generation. In the process it was gradually altered so that in the end it became a story about a shaman with magical powers. Those changes were accompanied by alterations in the name of Kwende's founder, which in the end became 'Mapemba'.

Assuming that the above story is true, answer the following question: when Thenga uses the name 'Mapemba', is he actually talking about the inventive leader Ndembo, who is the original source of the Mapemba legend, or is he talking about a fictional person, someone who does not really exist [the Fictional Person version] / or is he talking about no one in particular [the No One In Particular version] / or maybe he is not referring to anyone [the No Reference version]?

HOMER

*Jacques is an inhabitant of 16th-century France. Like most of his well-educated contemporaries, Jacques believes that Homer, a nomadic blind poet living in the 7th century BC, was the author of *The Odyssey*, a famous ancient Greek epic. Jacques acquired this belief while studying at leading medieval universities in Europe. But the truth is different.*

*The *Odyssey* is a piece of work that has no single author. The inspirations for the story depicted in *The Odyssey* can be found in the tales told by Callicrates, a story-teller living in ancient Greece in the 10th century BC. Callicrates was neither blind nor did he travel much in his life. His stories*

were so popular that they spread around Greece and people passed them on from generation to generation.

As the time passed, the stories were altered, some elements were replaced by new ones, some of them disappeared. Many people contributed to the final version of *The Odyssey*. In the end, the story does not have much in common with Callicrates' original tales. Along with the changes of the content and of the form of the story, people were changing their beliefs about the author, his life and his name. They started to believe that *The Odyssey* – which is the title that appeared with all the other modifications – was written by a nomadic blind poet called Homer.

Assuming that the above story is true, answer the following question: when Jacques uses the name 'Homer' is he actually talking about the storyteller Callicrates, whose tales were the inspiration for *The Odyssey* and who is the original source of the Homer legend, or is he talking about a fictional person, someone who does not really exist [the *Fictional Person* version] / or is he talking about no one in particular [the *No One In Particular* version] / or maybe he is not referring to anyone [the *No Reference* version]?

EINSTEIN

James is a high-school student living in Tinsbury, a small town in the south of England. Like most of Tinsbury inhabitants who attended high-school in their hometown, James believes that Albert Einstein was a physicist who invented the atomic bomb. Like most of residents of Tinsbury, James hasn't got any other beliefs concerning Albert Einstein. The truth is different, however.

The atomic bomb was not invented by Albert Einstein. In fact it was not invented by any single person but by a large group of scientists who participated in the Manhattan Project in the USA during World War II. Among others, Robert Oppenheimer, Ernest Lawrence and Harold Urey were involved in this project. Albert Einstein, who is famous mostly for his contribution to the Theory of Relativity, had never worked on the atomic bomb.

James' belief concerning Einstein is due to a mistake of an aged Physics teacher. Thinking that the atomic bomb is one of the greatest inventions of the 20th century and believing that Albert Einstein was the most eminent physicist of that century, the teacher ascribed this discovery to Einstein by mistake.

Assuming that the above story is true, answer the following question: when James uses the name 'Albert Einstein' is he actually talking about Albert Einstein, the author of the Theory of Relativity, who was the source

of the teacher's mistake, or is he talking about a fictional person, someone who does not really exist [the Fictional Person version] / or is he talking about no one in particular [the No One In Particular version] / or maybe he is not referring to anyone [the No reference version]?

4.4. Subjects

The experiment took the form of an electronic survey published online. As has been noted earlier, the language of the study was English and the participants were native English speakers, mainly from the United States of America, but also from Great Britain. The respondents received an invitation to participate in the study with a hyperlink to the survey via email (in the email, they were also encouraged to pass the invitation on and ask others to participate).

The respondents were volunteers and did not receive any remuneration for their participation in the study. 136 persons responded to the survey; 22 submissions were eliminated from further analysis because the persons in question were not native speakers of English or because they reported having undergone philosophical education at the level of Bachelor's degree or higher. The data presented below are from a group of 114 respondents.

39 persons were assigned to the Fictional Person version, 38 to the No One In Particular version, and the remaining 37 persons, to the No Reference version. 56.1% of the sample were women, 43.9% were men. The youngest participant was 18, and the oldest 71 years old – the average age of the sample was 34.4 years old with the standard deviation of 12.2. The majority of the respondents, 62.5% to be precise, were persons not older than 35.

4.5. Results

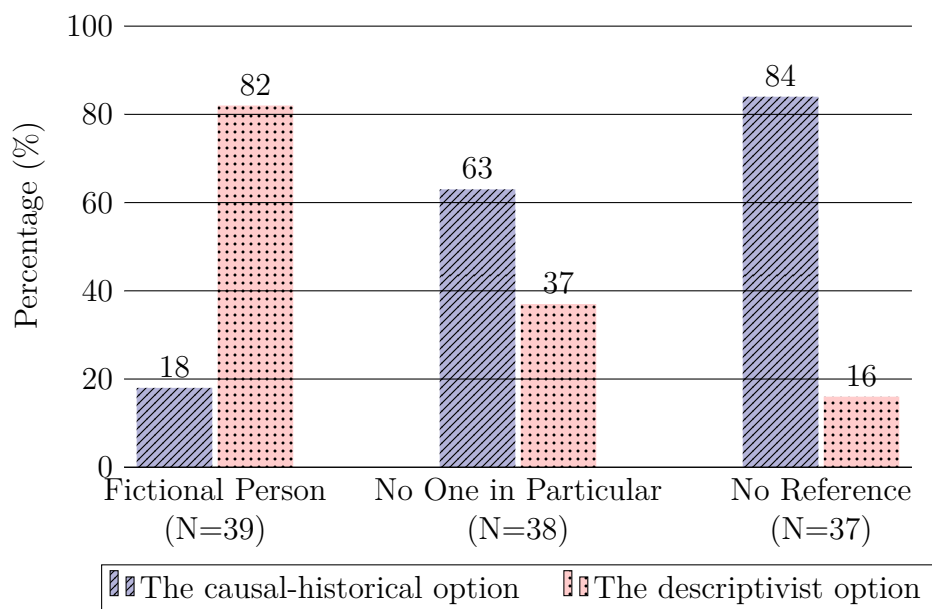
4.5.1. Comparison of the alternative formulations of the descriptivist option

I begin the presentation of the results by comparing respondent reactions to the scenarios depending on the particular version, that is, the formulation of one of the response options available to the respondents¹².

For the *Mapemba* scenario, clear and statistically significant differences in response distribution between the particular versions of the scenario

¹²Since the method used to gather the data (forced choice from several options, also referred to as closed-ended multiple choice) only allows for the measurement of the dependent variable on a nominal scale, the statistical tests used in the analysis were based on the comparison of frequency for the appropriate categories (response options). χ -square and Z tests were used.

were observed¹³. In the case of the *Fictional Person* version, a significant majority of respondents chose the option considered to express support for descriptivism in the original study by Machery *et al.* Namely, they concluded that the protagonist of the story referred to the fictional person who does not really exist. However, the situation was different in the case of the *No One In Particular* and *No reference* versions – persons expressing support for the causal-historical theory constituted the majority in each group. The percentage of responses supporting the causal-historical theory in these cases was significantly higher than in the case of the *Fictional Person* version¹⁴. One can thus speak of a “reversal” in respondent intuitions between the *Fictional Person* version on the one hand, and the *No One In Particular* and *No Reference* versions, on the other. The described result is illustrated on Graph 3.



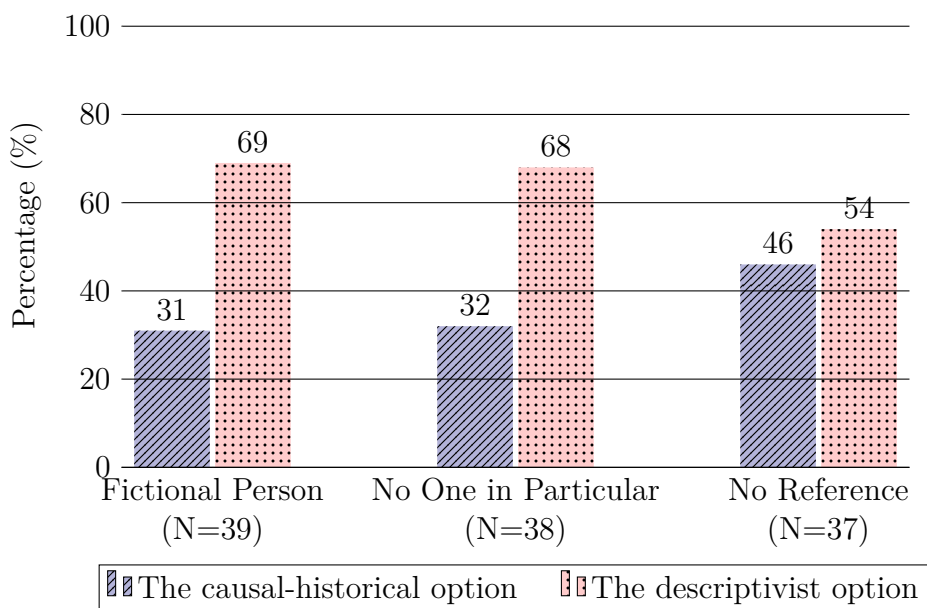
Graph 3. Response distribution for the *Mapemba* scenario depending on the version.

Unlike in the case of *Mapemba*, in the case of *Homer*, χ -square and Z tests did not indicate any significant differences in response distribution between the particular versions of the scenario. In the case of the *Fictional Person*

¹³ $\chi^2(2) = 34.94; p < 0.001$.

¹⁴Z tests (significance level $p = 0.05$).

and *No One In Particular* versions, persons selecting the option supporting the causal-historical theory constituted a clear minority, around 1/3 of the respondents. A slightly different situation occurred in the case of the *No Reference* version – here, the preferences of the participants were distributed almost equally, with slightly more than half choosing the descriptivist option stating reference failure. However, statistically adequate testing does not in fact permit the conclusion that this option was more popular than the competing one¹⁵. In the case of the other two versions, the predomination of the descriptivist options was statistically significant¹⁶. Therefore, despite the fact that the tests cited at the beginning of the paragraph did not show any significant differences in response distribution between the particular versions, it seems justified to speak of a clear tendency. The results for the *Homer* scenario are illustrated on Graph 4.



Graph 4. Response distribution for the *Homer* scenario depending on the version.

The formulation of the descriptivist option also had an impact on respondent preferences as far as the *Einstein* scenario is concerned¹⁷. In

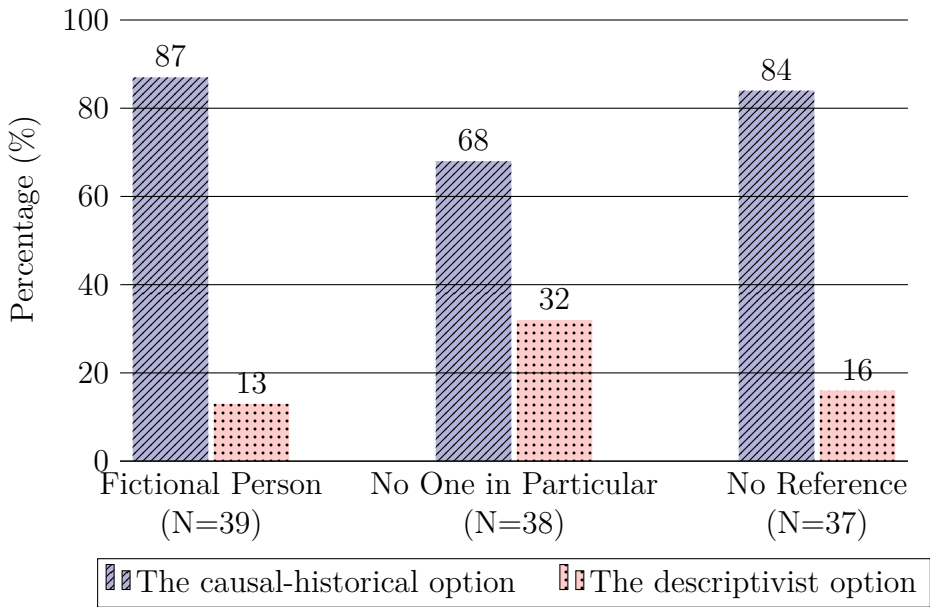
¹⁵ $\chi^2(1) = 0.24$; not significant.

¹⁶Fictional Person $\chi^2(1) = 5.77$; $p = 0.016$. No One In Particular: $X^2(1) = 5.16$; $p = 0.023$.

¹⁷ $\chi^2(2) = 4.94$; $p < 0.047$.

the case of the *Fictional Person* and the *No Reference* versions, responses supporting the causal-historical theory, that is, those stating reference to Albert Einstein, predominated decisively. However, in the case of the *No One in Particular* version, the predomination of the causal-historical option over the descriptivist one was slightly smaller – here, close to 1/3 of the respondents concluded that the protagonist did not refer to anyone in particular. The results of the experiment for the *Einstein* scenario are illustrated on Graph 5.

To summarize, the results of the study in terms of the comparison of the alternative formulations of the descriptivist option permit the conclusion that the impact of this factor was observed for all three tested scenarios, although in some cases it was weaker than in others.

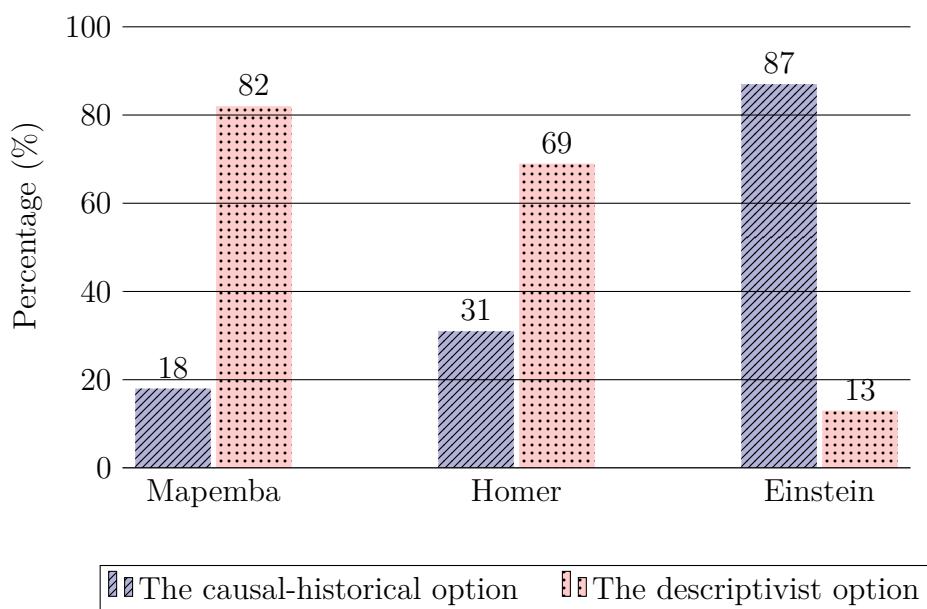


Graph 5. Response distribution for the *Einstein* scenario depending on the version.

4.5.2. Comparison of respondent reactions depending on the particular Jonah type scenario

The second goal of my first methodological study concerning the reference of proper names was to determine if respondent preferences for the analyzed conceptions of reference would be stable across different Jonah case scenarios.

For the formulation stemming from the original study by MMNS, that is, the one including the option that the protagonist of the story referred to a fictional person, significant differences in response distribution were observed between *Einstein* on the one hand, and *Mapemba* and *Homer*, on the other. In the first case, the majority of respondents concluded that the user of the name “Einstein” in fact referred to Albert Einstein; in *Mapemba* and *Homer*, the majority opined that the protagonist referred to a fictional person¹⁸. It thus turns out that even in the case of the standard approach to Jonah cases proposed by MMNS the preferences of nonphilosophers are not stable across different scenarios of this type. This result is summarized on Graph 6.

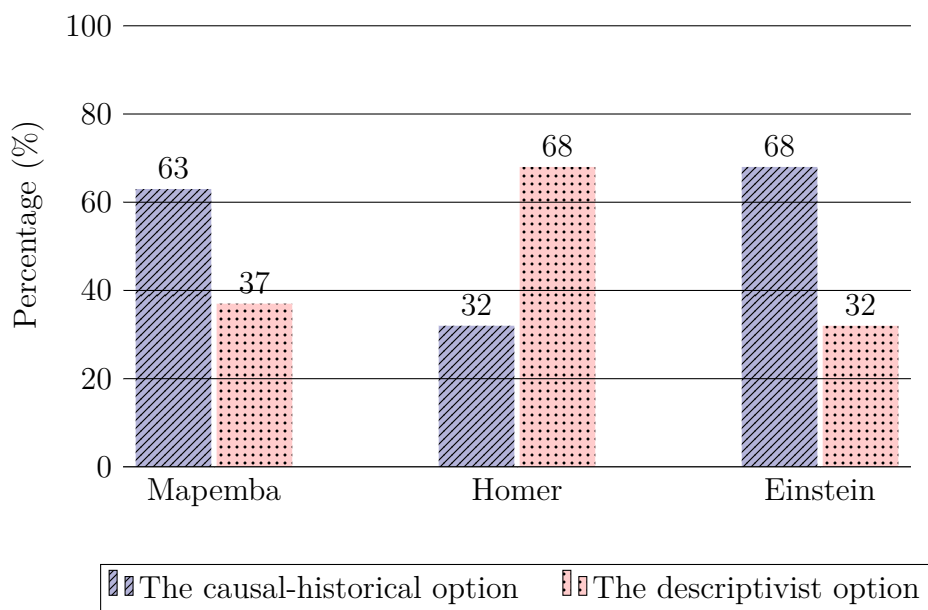


Graph 6. Response distribution for the particular scenarios in the case of the *Fictional Person* version (N = 39).

Similarly significant, albeit differently distributed, differences were observed in the case of the *No One In Particular* version. The respondents

¹⁸ $\chi^2(2) = 42.7$; $p < 0.001$ and adequate comparisons using the Z test indicated that these differences are statistically significant.

assigned to this version assessed the *Mapemba* and *Einstein* scenarios in a very similar manner – here, responses supporting the causal-historical theory predominated. However, in response to *Homer*, the majority of respondents stated that the protagonist did not speak of anyone in particular¹⁹. This result is shown on Graph 7.



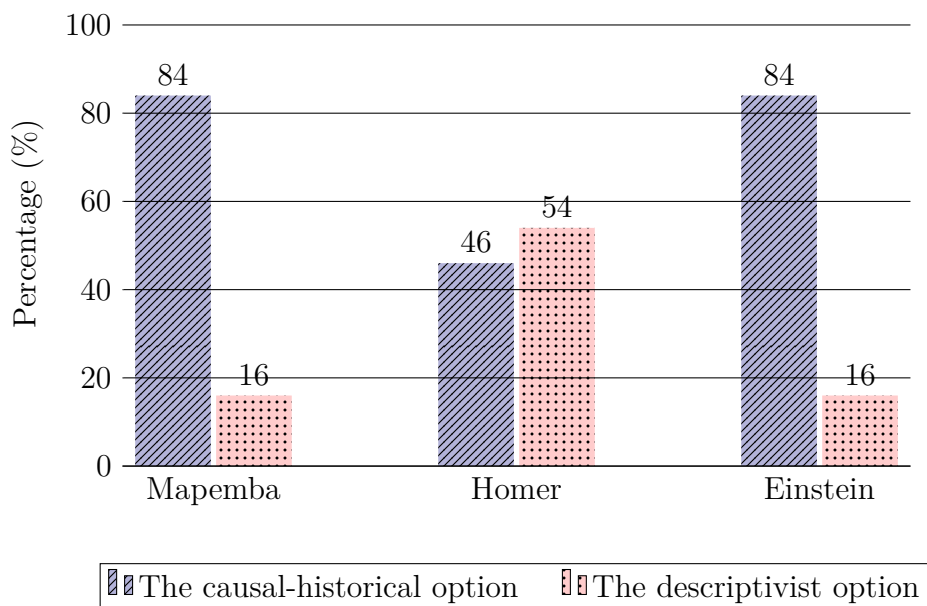
Graph 7. Response distribution for the particular scenarios in the case of the *No One in Particular* version (N = 38).

A clear divergence in respondent reactions to the particular scenarios also occurred in the case of the *No Reference* version. Just like in the case of *No One in Particular*, here too, respondent reactions to *Mapemba* and *Einstein* were very similar – in both cases the response stating reference failure was highly unpopular. In the case of *Homer*, on the other hand, none of the options predominated – respondent preferences were distributed almost evenly between the descriptivist and the causal-historical options²⁰.

¹⁹justifying The statistical significance of these differences has been confirmed by both the χ^2 -square test, $\chi^2(2) = 12.16$; $p < 0.002$, and appropriate comparisons using the Z test.

²⁰Differences in respondent reactions to *Homer* relative to *Mapemba* and *Einstein* turned out to be statistically significant according to both the Z tests ($p = 0.05$) and the χ^2 -square test $\chi^2(2) = 17.21$; $p < 0.001$.

Detailed information regarding the response distribution for this version are shown on Graph 8.



Graph 8. Response distribution for the particular scenarios in the case of the *No Reference* version (N = 37).

A summary of the comparison of the results for the particular scenarios should emphasize the fact that in the case of each formulation of the descriptivist option, the respondents reacted to one of the scenarios differently than to the other two. Interestingly, it was not the same scenario – in the case of the *No One in Particular* and *No Reference* versions, the outlying scenario was *Homer*; in the case of the original formulation from the study by Machery *et al.*, it was *Einstein*. It is worth noting at this point that the picture implied by the data gathered during my experiment is far more complex than has been the case in previous research dedicated to the problem of the reference of proper names. However, before I proceed to interpret these results, I would like to discuss the results of an additional study I conducted since they can help cast more light on the data presented so far.

4.6. The additional experiment

4.6.1. General characterization

There are several significant differences between the Jonah type scenarios used in my first experiment. I address them later in the paper. However, one difference between the *Einstein* scenario and the *Mapemba–Homer* pair is so fundamental that it calls for separate treatment. Namely, in the case of *Einstein*, the proper name borne by the referent intended by the protagonist and the name used by the protagonist have the same shape; the protagonists of *Mapemba* and *Homer*, in contrast, use names different in shape from those borne by the possible referents of these names (“Ndembo” vs. “Mapemba” and “Callicrates” and “Homer”). As is implied by the results of the first experiment, respondent reactions to the *Einstein* scenario were only slightly sensitive to the formulation of the descriptivist option – responses stating that the protagonist actually referred to Albert Einstein predominated in the case of each version. As regards *Mapemba* and *Homer*, the descriptivist response predominated in the case of at least one version. Perhaps this difference in the shape of the name had a significant impact on respondent reactions? The main goal of my second study was to empirically test this supposition.

In light of the above, alternative *Mapemba* and *Homer* scenarios were used in the second study in which the names of the persons whose actions inspired the legends responsible for the false beliefs of the described linguistic communities had the same shape as the names used by the protagonists. This was the only difference between the original scenarios and those analyzed in the additional study.

Each participant in the second experiment learnt two scenarios – *Mapemba*’ and *Homer*’ – presented in random order. The respondents were randomly assigned to one of the three versions of each scenario: *Fictional Person*, *No One in Particular*, or *No Reference*. The characteristics of these versions were analogous to the first study – the versions only differed in the formulation of the descriptivist option.

4.6.2. Subjects

Just like the initial study, the second survey was carried out over the Internet. The participants were recruited via the online portal Amazon Mechanical Turk (www.mturk.com) – registered users had access to a hyperlink to the survey and could commit to taking it in exchange for a fee of 0.30 USD. This time, then, the participants were not volunteers. 156 persons

filled out the survey; 21 respondents were not taken into account in later analysis because they reported having an academic degree in philosophy or were not native speakers of English. The statistics presented below are based on the sample size of 135 persons.

55.6% of the respondents were men, 44.4% were women. The youngest participant was 19 and the oldest one was 73 years old, the average age being 37.4 years old, with the standard deviation of 12.2. 60% of respondents were not older than 36. The distribution was thus slightly skewed toward younger persons.

4.6.3. Results

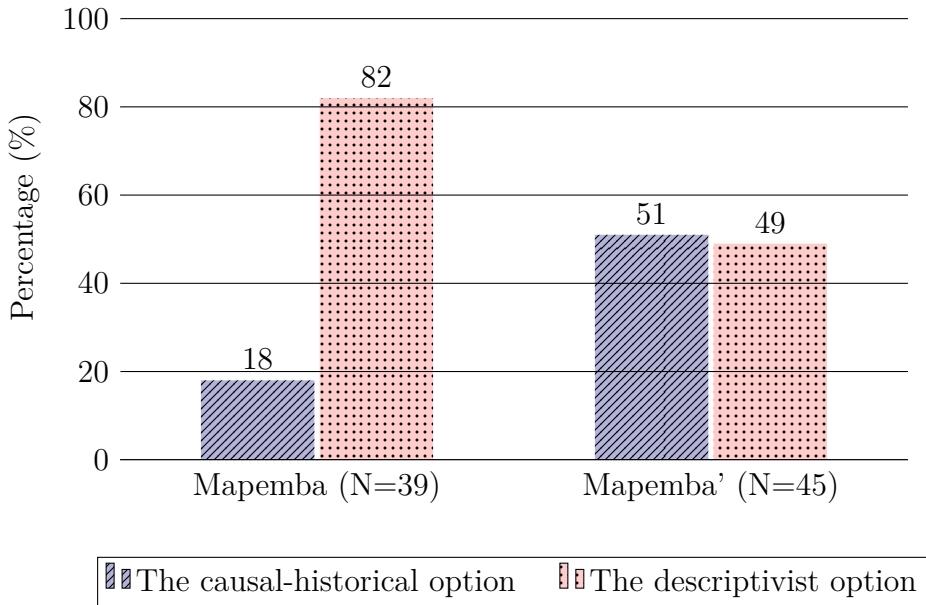
4.6.3.1. Respondent reactions to *Mapemba* and *Mapemba'* depending on the particular version

As part of the following statistical analysis, I compare respondent reactions to the *Mapemba* scenario (where the name used by the protagonist differs in shape from the name borne by the person who inspired the legend, here called Ndembo) and the alternative *Mapemba'* scenario (where the name used by the protagonist has the same shape as the name borne by the person constituting the possible referent of that name). A separate comparison was carried out for each version, that is, for each formulation of the descriptivist option.

Graph 9 illustrates the response distribution for the alternative scenarios in the case of the *Fictional Person* version. Let us remember that in the case of the original scenario responses stating reference to a fictional person predominated. In the case of the alternative scenario, where the leader was in fact called *Mpemba*, in contrast, the response stating reference to a fictional person was almost as popular as the one pointing to the ingenious leader as the referent. Differences in response distribution for the alternative scenarios are statistically significant²¹.

Sameness of shape of the name used by the protagonist and the name borne by the person who inspired later false beliefs shared by the protagonist had a similar impact in the case of the *No One in Particular* version. Although here responses supporting the causal-historical theory predominated both for *Mapemba* and *Mapemba'*, in the first case this predomination was clearly smaller than in the latter case, where the two names had the same

²¹ $\chi^2(1) = 10.01; p = 0.002$.



Graph 9. Response distribution for *Mapemba* and *Mapemba'* in the case of the *Fictional Person* version.

shape. Just like in the case of the *Fictional Person* version, here too, the difference is statistically significant²². The result is illustrated on Graph 10.

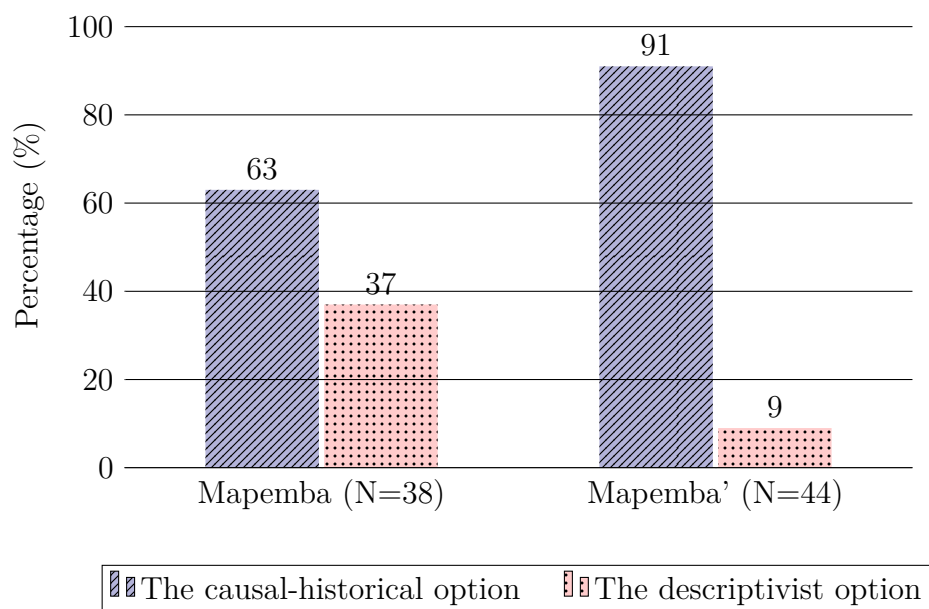
At the same time, no divergence in respondent preferences occurred in the case of the *No Reference* version²³. Here, a clear majority of respondents concluded that the protagonist of the story referred to the ingenious leader whose achievements inspired legend regardless of whether the shape of the name used by the protagonist and the shape of the name borne by the leader were the same or not. This result is shown on Graph 11.

4.6.3.2 Respondent reactions to *Homer* and *Homer'* depending on the particular version

The distribution of respondent preferences in regard to *Homer'*, where the name borne by the person who inspired later generations of poets to create the *Odyssey* was in fact *Homer*, in the case of the particular versions was similar to that observed for *Mapemba'*.

²² $\chi^2(2) = 9.17; p < 0.002$.

²³ $\chi^2(1) = 0.02$; not significant.



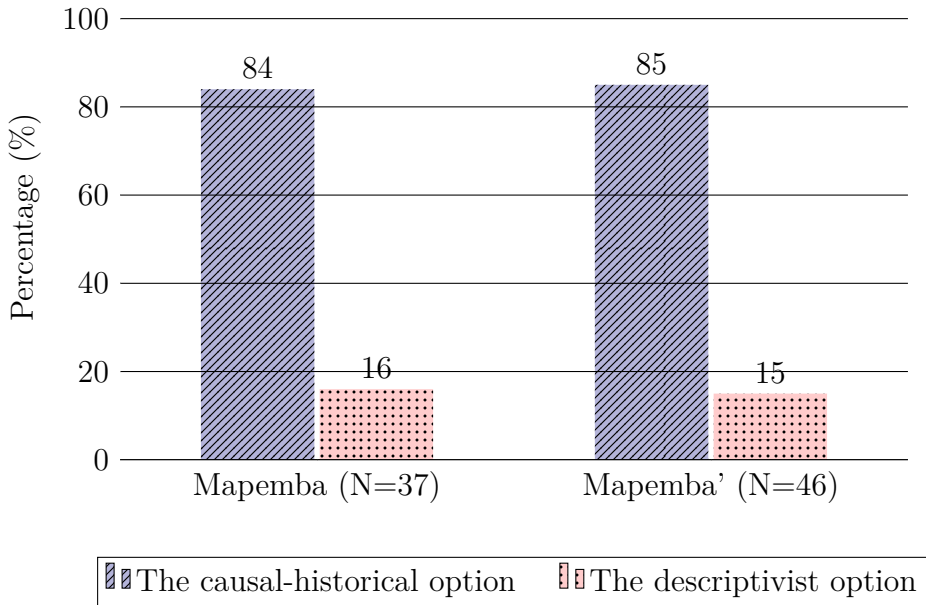
Graph 10. Response distribution for *Mapemba* and *Mapemba'* in the case of the *No One in Particular* version.

In the *Fictional Person* version, participant responses were divided almost in half. To recall, the original scenario, where the story teller who inspired the creation of the *Odyssey* was called Callicrates, led the majority of respondents to conclude that the protagonist referred to a fictional person. Differences between the two alternative scenarios were at the level of a statistical trend²⁴. The relevant result is shown on Graph 12.

Graph 13 shows the distribution of results for the *No One in Particular* version. Here, a slight difference between *Homer* and *Homer'* led to the complete reversal of respondent preferences. In response to the original articulation, where the names differed in shape, the majority of respondents concluded that the user did not refer to anyone in particular. The altered scenario, where the story teller who inspired the creation of the *Odyssey* was in fact called *Homer*, in contrast, led the majority of respondents to conclude that the protagonist referred to this person. These differences turned out to be statistically significant²⁵.

²⁴ $\chi^2(1) = 3.56; p = 0.059$.

²⁵ $\chi^2(1) = 17.29; p = 0.001$.

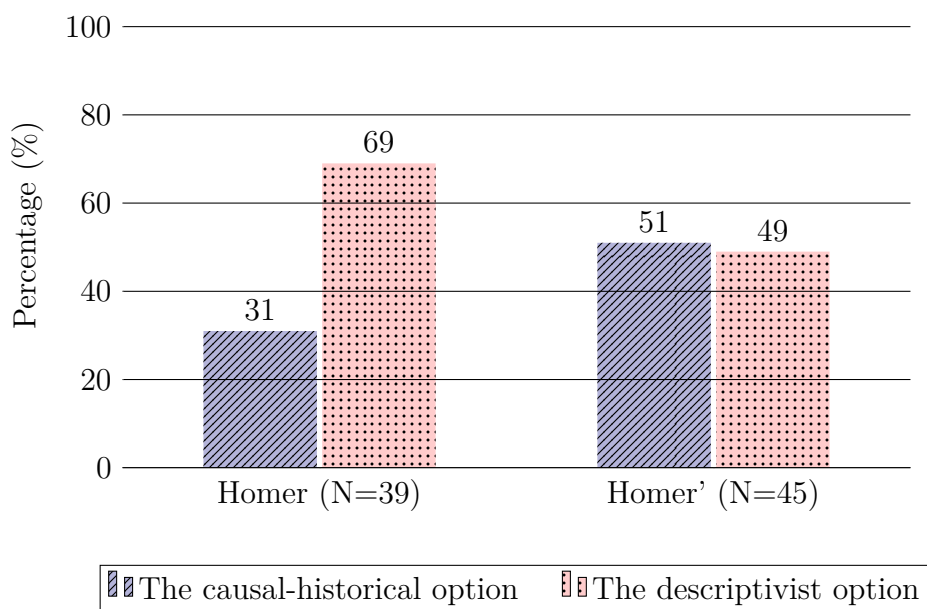


Graph 11. Response distribution for *Mapemba* and *Mapemba'* in the case of the *No Reference* version.

The change in *Homer'* relative to *Homer* also affected participant responses in the case of the *No Reference* version – a statistically significant divergence depending on the shape of the name of the person constituting the possible referent was observed²⁶. When this person bore the name “Callicrates,” no response predominated (slightly more than a half of respondents concluded that the protagonist did not refer to anyone); however, when his name had the same shape as that used by the protagonist, the opinion that the protagonist referred to that person clearly predominated. This result is shown on Graph 14.

To summarize the results of my second methodological experiment, one could say that, according to predictions, sameness of shape of the name used by the protagonist and that belonging to the person who inspired legend led to a significant increase in the percentage of responses supporting the causal-historical theory relative to the scenarios used in the first experiment, where these names differed in shape. In the case of *Homer* this tendency was observed for all three versions. In the case of *Mapemba*, it was observed for the *Fictional Person* and *No One in Particular* versions. It should perhaps

²⁶ $\chi^2(1) = 10.73; p = 0.001$.



Graph 12. Response distribution for *Homer* and *Homer'* in the case of the *Fictional Person* version.

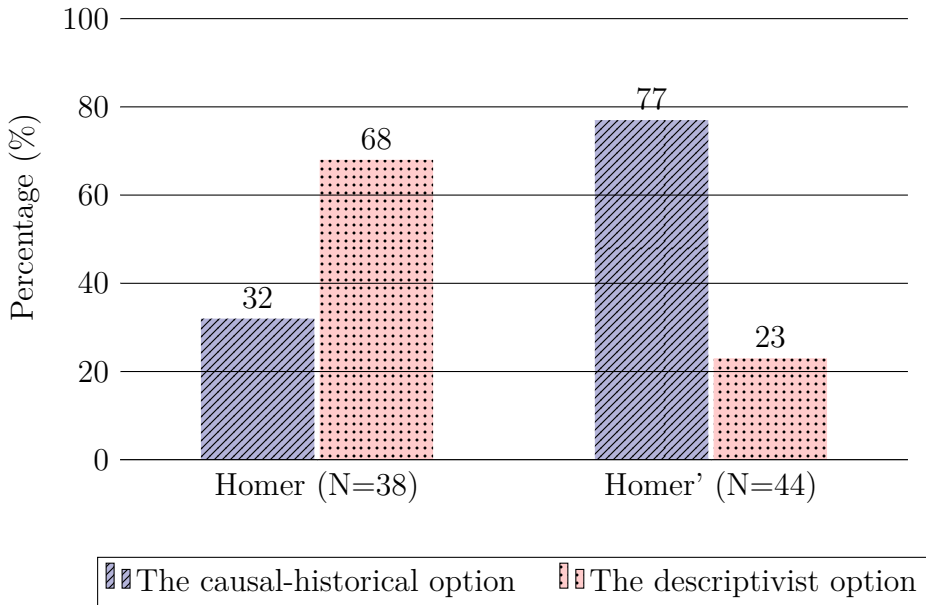
be emphasized that in some cases (*Homer*, the *No One in Particular* version) the manipulation of the content of the scenario had such an immense impact on participant responses that it led to the complete reversal of preferences.

5. Results

The picture emerging from the data gathered during my studies is fairly complex, especially in light of the results obtained in previous research concerning the reference of proper names. In what follows, I present two different interpretive strategies leading to different conclusions. At the same time, I should point out that these two approaches certainly do not exhaust the set of all consistent explanations of the obtained data. I should also stress that the discussed results ought to be approached cautiously since the conclusion is based on limited empirical material.

5.1 The first interpretive strategy: the data support the position of Machery *et al.*

One possible interpretation of the data presented here is that they in fact support the main thesis put forth by Machery *et al.* (2004). To

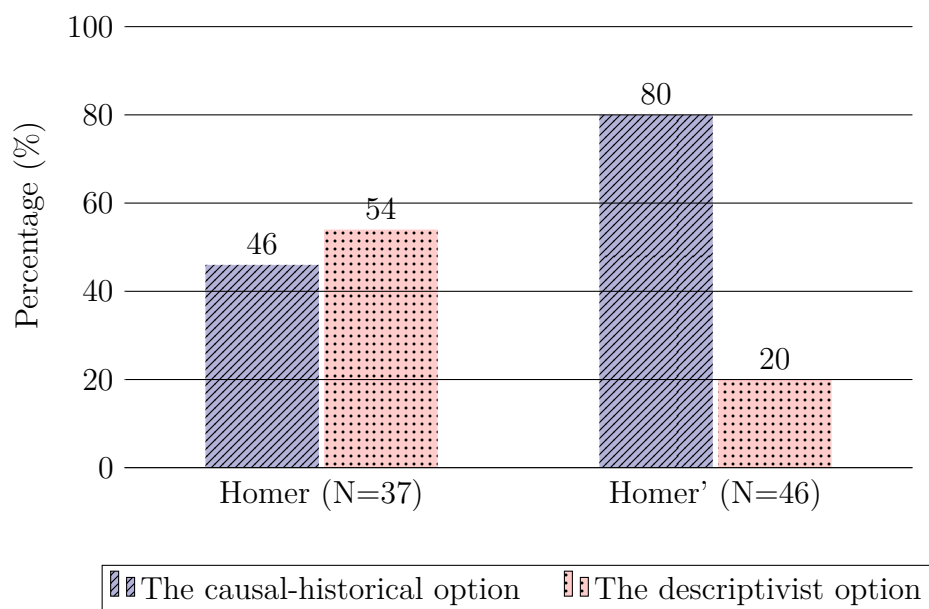


Graph 13. Response distribution for *Homer* and *Homer'* in the case of the *No One in Particular* version.

recall, MMNS claim that intuitions about the reference of proper names are susceptible to the impact of philosophically insignificant factors. In the case of their research, this factor was the cultural background of persons expressing opinions regarding reference. My research is not cross-cultural and thus does not constitute further evidence regarding the impact of this factor on semantic intuitions. However, the results of my study, targeting predominantly members of one nation (citizens of the USA), demonstrate an intracultural and even intrapersonal variation in intuitions about the reference of proper names. This variation seems to be due to factors that should not impact such intuitions²⁷. In particular, no divergence should have occurred between respondent preferences in regard to the different Jonah type scenarios used in the study since, as far as philosophically significant aspects are concerned, these scenarios are similar. Such divergence was nonetheless noted.

Before I develop and assess the idea that the obtained data could be interpreted as supporting MMNS, I must point to one aspect that constitutes

²⁷For more detail regarding these factors and their impact, see Section 5.2. dedicated to the discussion of the second interpretive strategy.



Graph 14. Response distribution for *Homer* and *Homer'* in the case of the *No Reference* version.

a serious problem for them. Namely, it turned out, according to predictions, that changes in the formulation of the descriptivist option relative to the formulation used in the original MMNS study led to a variety of different participant responses. It is thus not possible to claim that by choosing alternative formulations of the descriptivist option the respondents expressed the same intuitions – not every formulation permitted the respondents to express their support for descriptivism. As I have argued earlier, there are independent reasons to maintain that the option stating reference to a fictional person used in the study by MMNS is in fact inadequate. Therefore, it is justified to claim that their experiment – at least in the part focusing on Jonah cases – did not measure the preference of nonphilosophers for the causal-historical *versus* the descriptivist theory but their preference for the first versus some other conception of the reference of proper names (one allowing for the possibility of referring to a fictional object).

Moreover, even assuming that I have used the right operationalization of descriptivism in Jonah type situations is either the *No One in Particular* or the *No Reference* version, respondent preferences in these cases are still ambiguous and unstable across the tested scenarios. In the *No Reference*

version, where the formulation of the descriptivist option is perhaps closest to the spirit of descriptivism, preferences for the causal-historical theory clearly predominated in the case of *Mapemba* and *Einstein*; at the same time, no response predominated in the case of Homer. Therefore, if this is considered to be the right approach to determining which responses support the descriptivist theory and which accord with the causal-historical conception in Jonah cases, it does not eliminate the instability characterizing the responses of nonphilosophers – not so much on the cross-cultural level as within one culture and even between the individual assessments made by the same person. Hence, the dose of suspicion MMNS propose to apply to semantic intuitions seems even more justified in light of my results.

The results of my research can thus be read as supporting the main thesis put forth by Machery *et al.*: that semantic intuitions about the reference of proper names are unstable and uncertain and therefore useless in philosophical debates. However, let us note that their conclusion is based on a hitherto unquestioned assumption that, in philosophical-experimental research of the type similar to theirs, by choosing one of the available response options the respondents do in fact express support for certain theories of reference. Their explanation for the cross-cultural differences observed in much research is consistent with this assumption. The reason for an increased tendency among the Chinese, relative to the Americans, to express intuitions supporting the descriptivist theory is supposed to be a tendency prevalent among the latter (and characteristic of all members of Western culture) to perceive reality in causal terms. The preference for the causal-historical theory among members of Western culture is thus presumably an effect of a more general phenomenon, namely, a preference for a particular cognitive strategy characteristic of that culture.

The key feature of the argument presented by Machery *et al.* is the observation that the majority of analytic philosophers, who tend to support the causal-historical conception, are persons raised in Western culture. Since according to MMNS there is no basis for the claim that philosophers are not susceptible to the influence of culture, it is safe to assume that their preferences too are culturally conditioned, especially since the majority have been raised in the Western cultural sphere. And here lies the crucial problem – the results of my experiments (and research results showing relevant differences within Western culture in general) do not warrant such an easy and fluid passage from tendencies observed in nonphilosophers to

alleged tendencies in the semantic intuitions of trained philosophers²⁸. To wit, I observed clear intrapersonal differences in the responses of nonphilosophers to different Jonah type scenarios. I am of the opinion that a similar effect would not occur in philosophers – the commitment to avoid contradiction and to maintain the consistency of expressed views is one of the top priorities of the philosophical academy (at least in so-called analytic philosophy). Moreover, it does not seem plausible that the impact of factors affecting participant responses observed in my experiments could be explained in a manner tracking one of the discussed theories of reference as closely as the cross-cultural explanation proposed by MMNS. Let us consider the strong impact of the shape of the name used by the protagonist on participant responses observed in my experiment. Whereas it can be expected that the tendency to support the causal-historical theory of the reference of proper names should correlate positively with the tendency to perceive reality in causal terms, the impact of the shape of the name on preferences for either the causal-historical or the descriptivist conception is unexpected. As we have seen, sameness of shape of the name used by the protagonist and that borne by the person constituting the possible referent of that name translates into a significant increase in the percentage of responses interpreted as supporting the causal-historical theory. Meanwhile, this factor should not significantly impact preferences for this theory since a change in the shape of the name does not change the causal chain linking its use to the referent (if such an object exist) – the most important feature of that conception. The case of descriptivism is similar. Here, the description tied to the name is crucial since it is the description that determines the referent; the shape of the name does not play any role. The impact of the shape of the name on preferences guiding the selection of either the causal-historical or the descriptivist theory is thus difficult to explain. A doubt therefore arises: do the participants in experiments based on methodologies similar to that used by MMNS (the choice of determinate response options) in fact express support for either descriptivism or the causal-historical theory, if only in a trivialized version reconstructed by MMNS?

If the above line of reasoning is correct, then another strategy of interpreting the data gathered during my research must be pursued. According

²⁸To recall, another problem connected to this explanation is the fact that its correctness is doubtful. As has turned out in the course of the experiment by Machery *et al.* (2009), the French, that is, members of Western culture, supported the causal-historical theory in the case of Gödel situations less willingly than did Mongols, members of Asian culture.

to the alternative interpretation presented below, methodologies modelled on the approach adopted in the study by Machery *et al.* (2004 and 2009) are not effective as far as revealing the semantic intuitions of nonphilosophers is concerned. This is because many participants in studies of this kind express intuitions that are not semantic.

5.2 The second interpretive strategy: the data do not reflect semantic intuitions

According to the second interpretive approach, the data gathered during my methodological experiments concerning the reference of proper names can be used to undermine the key assumption shared by the authors of research to date: that by selecting a particular response in Jonah type situations the respondents in fact express their support for either descriptivism or the causal-historical theory (this conclusion could perhaps be generalized to include Gödel cases; this would require further empirical research). The reason for this interpretation is the instability of participant responses observed in the course of my experiments. In what follows, I explain in detail which aspects of the scenarios used in my studies translated into differences in respondent preferences and what their impact might consist in. However, before I proceed to this detailed interpretation, I would like to characterize the general mechanism I consider to be responsible for participant reactions to experiments based on methodologies modelled on the study by MMNS.

According to the second interpretive strategy, at least some of the persons participating in my studies selected the responses they did not due to their semantic intuitions about the reference of proper names but due to heuristics based on simple associations: they focused on certain “superficial” verbal tips embedded in the individual scenarios. Both Gödel and Jonah cases are complex and exceptional situations manageable by philosophers but not by persons lacking philosophical training. It can thus be expected that the semantic intuitions of the latter are not well grounded and simply break down when such uncommon situations are considered, giving no basis for unambiguous and certain responses. It is more than likely that many nonphilosophers, when confronted with stories they found troublesome and lacking clear guidance stemming from their linguistic competency, tried to imbue these stories with some sense by associating them with other scenarios, previously encountered in daily life, and by adopting simplified coping strategies.

There are many reasons for claiming that the responses of at least some respondents participating in my experiments resulted from superficial

information processing and simplified techniques of analysis. In my opinion, almost all systematic differences in respondent preferences between the particular Jonah scenarios and the particular versions of these scenarios are of this sort. This is especially striking once the most salient and commonsensical explanation of these differences is taken into account.

Let us consider my first experiment and differences in participant responses to each scenario in the *Fictional Person* version. How can we explain, in simple terms, the stark difference in participant responses between *Einstein* on the one hand, and *Mapemba* and *Homer*, on the other? The construal of the content of *Mapemba* and *Homer* activated in the respondents a cognitive schema for processing fictional and legendary persons. The description tied to the name in the first story mentioned the possession of magical powers, a feature considered by most to be fictional in and of itself; the second story referenced the hypothesis still being explored in the history of literature that Greek epics were not created by single author and that Homer is a legendary figure (many of the respondents may have learnt about this hypothesis in the course of their education). The Einstein scenario, in contrast, did not contain such elements (here, an erroneous belief of the linguistic community did not result from a long process of collective legend making but a single person's mistake) – the potential referent here is a person known to the majority of the respondents (including from photographs), a person whose existence was thus most likely not in doubt. The impact of this element was in fact so strong that the response stating reference to Albert Einstein predominated regardless of the formulation of the competing response (an effect which did not occur for the other two scenarios). It is thus possible that the basis for the respondents' reactions to the three scenarios consisted in superficial associations triggered by these scenarios. As my second experiment demonstrated, another factor responsible for significant differences in responses to these scenarios (albeit not for their entirety) was the fact that in *Mapemba* and *Homer*, but not in *Einstein*, the name used by the protagonist differed in shape from the name borne by the possible referent. Here too, it seems justified to conclude that the impact of this manipulation was based on a simple mechanism of association.

Respondent preferences in the case of the *No One in Particular* version can be explained in a similar manner. The response option stating reference to no one in particular was preferred by the majority of respondents for *Homer*; in the case of *Einstein* a minority preferred it, although it was still the most popular of all three formulations of the descriptivist option. This is likely related to the fact that both these scenarios, unlike *Mapemba*, clearly

indicated that the description tied to the key name by the protagonist can be truthfully predicated of all members of a group even if it did not unequivocally point to any one individual. It seems that an association between the relevant situation and linguistic practices involving the expression “no one in particular,” especially pronounced in the English language, was decisive here.

The marginal popularity of the descriptivist option in the No Reference version is not surprising in this context either. If the majority of respondents perceived the scenarios in light of typical situations – situations featuring successful communication between interlocutors – it is possible that they rejected the response stating reference failure “in advance,” without further analysis, regardless of the user’s explicit intention. This preference may have been additionally strengthened by a familiar pragmatic phenomenon consisting in the tendency on the part of the recipient of an utterance to interpret it so that it possesses the desired semantic value (most often truth) – this can be linked to such philosophical concepts as Lewis’s principle of accommodation (e.g. Lewis, 1979) and Davidson’s principle of charity (e.g. Davidson, 1973).

On the other hand, one must admit that the response distribution for the *Homer* scenario in the *No Reference* version turned out to be slightly surprising. Here, the response stating reference failure was as popular as the one indicating reference to the story teller who inspired the creation of the *Odyssey*. However, the *Homer* scenario differs from the other two stories in a way which could provide a basis for the observed differences in respondent reactions. The specificity of the *Einstein* scenario relative to *Homer* (and *Mapemba*) consists in the fact that respondent reactions (the majority of respondents supported the causal-historical theory for all three versions of the scenario) could be shaped, to a large degree, by the conviction that all uses of the name “Einstein” refer to Albert Einstein – it is possible that they did not see any other option as attractive. The difference between *Homer* and *Mapemba*, on the other hand, could stem from the fact that whereas the protagonist of *Mapemba* shares certain adequate beliefs concerning the possible referent of the name he uses, the protagonist of *Homer* does not share any such beliefs. According to the latter scenario, Jacques believes that Homer is the author of the *Odyssey* and he does not have any other beliefs regarding this person – the description is false since, as it turns out, the *Odyssey* simply does not have an individual author. The protagonist of the *Mapemba* scenario, in contrast, not only believes that Mapemba was a shaman imbued with magical powers but also that

he was the founder of Kwende. As it turns out, even though the founder of Kwende was not a magic yielding shaman and his name was not Mapemba, he nonetheless existed. It is thus possible that the respondents assumed that the protagonist of the story successfully referred to the leader Ndembo since one description associated by him with the name “Mapemba” – the description “the founder of Kwende” – does identify Ndembo. The situation in *Homer* is different; here, the protagonist cannot be considered to share any belief correctly identifying Callicrates. It must be stressed at this point that this explanation of the differences between *Homer* and *Mapemba* points to questions of substance, relevant from the semantic perspective – it is thus not the sought-after pragmatic explanation of superficial differences between the scenarios. Despite this minor exception, the explanation of the observed differences presented here provides strong evidence for the conclusion that a crucial role in shaping participant responses in my experiments was played not so much (or not entirely) by the respondents’ semantic intuitions as by (broadly construed) pragmatic phenomena.

I am of the opinion that there is another reason to trust the explanation according to which participant responses are shaped based on superficial associations. Namely, the method of measuring respondent opinions used in research to date makes it difficult for the respondents to express their semantic intuitions elicited by the particular scenarios. Both the participants of the pioneering study by MMNS (2004) and those of many later experiments dedicated to the same issue (including the participants of my experiments) were forced to choose between two response options. The differences between the particular versions, differing in the formulation of one of the options, observed in my experiments demonstrate that many alternative and potentially attractive responses to Jonah scenarios can be indicated. The measurement of semantic intuitions, as this is the intention here – based on forced choice of one of two options limits the free expression of respondent opinion²⁹. On the one hand, this fact could be demotivating, discouraging the respondents from an in-depth consideration of the presented problems. As I have noted earlier, considering problems such as Jonah type situations, lying outside the scope of non-philosophical experience, is demanding enough – limiting the freedom of expression certainly made the task even more daunting. On the other hand, if in the opinion of the respondent none of the proposed response options correctly characterized the reference of the name in the described

²⁹I received several emails from the participants of my experiments registering the complaint that none of the proposed options reflected their intuitions, making it impossible for them to express their actual intuitions.

situation, but the respondent was forced to choose one option, it should be expected that they sought some kind of justification for their choice. But this justification need not be related to the respondent's semantic intuitions; it could refer to some other beliefs shared by the respondent or to aspects of the scenarios that are not necessarily relevant to the reference of proper names. Some candidates for such justifications have been indicated in the above attempt to explain the source of the differences observed in the course of my experiments.

It should be emphasized once more that my studies were methodologically similar to earlier research dedicated to the problem of the reference of proper names. The impact on respondent preferences was achieved by introducing minor differences into the content of the presented scenarios or the formulation of the response options available to the respondent – differences that should not influence, at least in theory and certainly not as strongly, semantic intuitions. Therefore, I am of the opinion that the key conclusion I have argued for could be generalized onto other experiments aimed at measuring nonphilosophers' opinions concerning the reference of proper names. Most probably not all participants of the studies carried out by Machery *et al.* (2004 and 2009) based their responses, including responses to Gödel cases, on their semantic intuitions, at least because the measuring tool employed in these studies was not sensitive enough to allow them to fully express such intuitions. Even on the assumption that, despite these difficulties, each respondent made sure that their responses are in fact close to their semantic intuitions, the interpretation of these choices as supporting either descriptivism or the causal-historical theory is still uncalled for. The spectrum of possible and philosophically consistent responses to Gödel and Jonah cases is undoubtedly far broader than the small range of options proposed by Machery *et al.* Forcing the respondents to choose one of them is unlikely to tell us anything precise about the semantic intuitions of nonphilosophers.

Conclusion

I am of the opinion that the data gathered during my studies and their discussion presented above constitute a relatively strong argument for doubting the methodology employed in research concerning the reference of proper names proposed by MMNS. In particular, the discussion demonstrates that more caution is needed in regard to the assumption, prevalent among

experimental philosophers, that responses to their surveys can be directly linked to philosophical intuitions of a certain kind.

The acceptance of the above argumentation should perhaps lead to the formulation of some additional methodological conclusions. Namely, since the structure of the studies dedicated to the problem of the reference of proper names carried out to date – one based on forced choice as the method of measuring semantic intuitions – limits the participants' freedom to express their opinions, a change in the way respondent preferences are measured should be considered. A natural alternative to forced choice is to give the respondents an opportunity to freely formulate their opinions or to choose from a broader range of options. The first solution is not popular among researchers – not only among experimental philosophers but in every social science reliant on survey methods, because data gathered this way is much more demanding and tedious to analyze than forced choice from a limited number of options. In order to carry out a quantitative analysis of such data it is necessary to group and categorize the free responses of the participants (coding). This task, if bias is to be avoided, should not be carried out by the authors of the experiment but by so-called competent judges – persons not familiar with the goals of the study and the hypotheses posed but competent enough to classify participant responses in a manner useful to the researchers. This further complicates research procedure. Another and perhaps more significant problem related to the analysis of free respondent feedback is the high percentage, noted for studies employing this method of measurement, of responses that are ambiguous, difficult to classify, or simply irrelevant to the question.

In light of the results of my experiments, formulating an open question which would encourage all participants to interpret the presented problem in the same way and to base their response on semantic intuitions (as opposed to some other kind) would be a particularly demanding task. Here, the use of a range of previously prepared responses can help curb ambiguity – in a sense, limiting the spectrum of responses makes the presented problem more determinate and precise, suggesting a perception desirable from the perspective of the experimenters. Therefore, a potentially interesting strategy could be to use surveys allowing the participants to freely express their opinion regarding the reference of proper names to determine a range of attractive and relevant responses which could be presented to respondents in multiple choice format in subsequent research. Such an approach would likely lead to obtaining relatively precise knowledge regarding the semantic intuitions of the respondents based on their reactions even to one scenario.

There is no doubt that further philosophical-experimental studies concerning the reference of proper names require an exploration of novel methodological avenues.

Bibliography

Alexander, J. (2012). *Experimental Philosophy. An Introduction*. Cambridge: Polity Press.

Bealer, G. (1998). Intuition and the Autonomy of Philosophy. In M. DePaul, W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* pp. 201–240. New York: Rowman & Littlefield.

Beebe, J. R., Undercoffer, R. J. (2016). Individual and Cross-Cultural Differences in Semantic Intuitions: New Experimental Findings. *Journal of Cognition and Culture*, 16(3–4), 322–357.

Cappelen, H. (2013). *Philosophy Without Intuitions*. Oxford: Oxford University Press.

Davidson, D., (1973). Radical Interpretation. *Dialectica*, 27, 314–328.

Deutsch, M. (2009). Experimental Philosophy and the Theory of Reference. *Mind and Language*, 24(4), 445–466.

Devitt, M. (2011). Experimental Semantics. *Philosophy and Phenomenological Research*, 82(2), 418–435.

Frege, G. (1997). Sens i znaczenie. In *Pisma semantyczne*, trans. B. Wolniewicz (pp. 60–88). Warszawa: Wydawnictwo Naukowe PWN.

Grice, H. P. (1975). Logic and Conversation. In P. Cole, J. Morgan (Eds.), *Syntax and Semantics*, vol. 3 (pp. 41–58). New York: Academic Press.

Jackman, H. (2009). Semantic Intuitions, Conceptual Analysis, and Cross-Cultural Variation. *Philosophical Studies*, 146(2), 159–177.

Kim, M., Yuan, Y. (2015). No Cross-Cultural Differences in the Gettier Car Case Intuition: A Replication Study of Weinberg et al. 2001, *Episteme*, 12(3), 355–361.

Knobe, J., Nichols, S. (2008). An Experimental Philosophy Manifesto. In J. Knobe, S., Nichols (Eds.), *Experimental Philosophy*, vol. 1 (pp. 3–14). Oxford: Oxford University Press.

- Kripke, S. (1972). Naming and Necessity. In D. Davidson, G. Harman (Eds.), *Semantics of Natural Language* pp. (253–355). Dordrecht: Reidel.
- Lam, B. (2010). Are Cantonese-Speakers Really Descriptivists? Revisiting cross-cultural semantics. *Cognition*, 115, 320–329.
- Lewis D. (1979). Scorekeeping in a Language Game. In: R. Bauerle, U. Egli, A. von Stechow (Eds.), *Semantics from a Different Point of View* (pp. 172–187). Berlin: Springer.
- Ludwig, K. (2007). The Epistemology of Thought Experiments: First Person versus Third Person Approaches. *Midwest Studies in Philosophy*, 31, 128–159.
- Ludwig, K. (2010). Intuitions and Relativity. *Philosophical Psychology*, 23(4), 427–445.
- Lycan, W. G. (1999). *Philosophy of Language: A Contemporary Introduction*. London: Routledge.
- Machery, E. (2012). Expertise and Intuitions about Reference. *Theoria*, 27(1), 37–54.
- Machery, E., Deutsch, M., Mallon, R., Nichols, S., Sytsma, J., Stich, S. (2010). Semantic Intuitions: Reply to Lam. *Cognition*, 117, 361–366.
- Machery, E., Deutsch, M., Sytsma, J. (2015). Speaker's Reference and Cross-Cultural Semantics. In A. Bianchi (Ed.), *On reference* (pp. 62–76). Oxford: Oxford University Press.
- Machery, E., Mallon, R., Nichols, S., Stich, S. (2004). Semantics, Cross-Cultural Style. *Cognition*, 92, B1–B12.
- Machery, E., Olivola, C. Y., De Blanc, M. (2009). Linguistic and Metalinguistic Intuitions in the Philosophy of Language. *Analysis*, 69(4), 689–694.
- Marti, G. (2009). Against Semantic Multi-Culturalism. *Analysis*, 69(1), 42–48.
- Muszyński, Z. (2000). *Komunikacja i znaczenie: semantyczny aspekt komunikacji*. Lublin: Wydawnictwo UMCS.
- Nadelhoffer T., Nahmias E. (2007). The Past and Future of Experimental Philosophy. *Philosophical Explorations*, 10, 123–149.
- Searle, J. (1958). Proper Names. *Mind*, 67, 166–173.
- Seyedsayamdost, H. (2015). On Normativity and Epistemic Intuitions: Failure of Replication. *Episteme*, 12(1), 95–116.

Sytsma, J., Livengood, J. (2011). A New Perspective Concerning Experiments on Semantic Intuitions. *Australasian Journal of Philosophy*, 89(2), 315–332.

Weinberg, J., Alexander, J. (2014). The Challenge of Sticking with Intuitions through Thick and Thin. In A. Booth, D. Rowbottom (Eds.), *Intuitions* (pp. 187–212). Oxford: Oxford University Press.

Weinberg, J. M., Nichols, S., Stich, S. (2001). Normativity and Epistemic Intuitions. *Philosophical Topics*, 29(1–2), 429–460.

Williamson, T. (2007). *The Philosophy of Philosophy*. Malden, MA: Blackwell.

Originally published as “Odniesienie nazw własnych, intuicje semantyczne i filozofia eksperymentalna” *Studia Semiotyczne* 30/2 (2016), 45–97. Translated by Joanna Guzowska.