UNIVERSITY OF TARTU

DEPARTMENT OF PHILOSOPHY

Kaimar Karu

Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?

MA Thesis

Supervisor: Prof. Dr. Daniel Cohnitz

Tartu, 2010

Karu, Kaimar
Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?

Table of Contents

1. Introduction

In my thesis, I set out to analyze how to overcome some of the issues of using intuitions in philosophical arguments, a topic which has recently been brought back to active discussions by *experimental philosophers* in connection with the theories of reference. I begin by setting the scene and describing different theories of reference in a short overview.

I will then discuss the experimentalist challenge of the fallibility of intuitions and the resulting unsuitability of using intuitions in philosophy. I will show which studies have been conducted by experimental philosophers to reach the conclusion that there is a significant variance in intuitions both between and within demographic groups, leading to the experimentalist challenge.

This challenge could be avoided if we can show that philosophical arguments (e.g. by Kripke regarding the theory of reference), targeted by experimental philosophers, do not rely on intuitions. Max Deutsch has recently argued for this position, and I will show that this position is mistaken.

Another response to the challenge would simply be to accept its conclusions and give up using intuitions in philosophy. I will argue that although intuitions are far from being infallible, they can still be useful and we should not give up using intuitions in philosophy. I will then answer the experimentalist challenge.

Several philosophers, also considering intuitions as a relevant source of knowledge, have claimed that it is only philosophers' intuitions that count, and that folk intuitions are irrelevant to philosophical debates. I argue that there is no basis for asserting philosophers' intuitions a higher status compared to folk intuitions in terms of their likeliness to be true, but they have a higher potential of giving us relevant information.

In a situation where philosophers can access their intuitions by engaging in philosophical contemplation, the methods for harvesting folk intuitions are still immature. I will present several types of methodological criticism towards the experimental philosophers' studies that were designed to elicit intuitions from the folk participants. At the same time, I will list my recommendations for improving the current methodology.

The penultimate chapter is dedicated to subsequent studies in experimental philosophy. These show the recent trends where experimental philosophy is heading, but also in some cases already the results of applying methodological improvements to the experiments.

In the last chapter I discuss the paths experimental philosophy ought to take to retain (or achieve, if a sceptical position is currently held) philosophical significance and provide us with relevant data.

2.   Theories of reference

Theories of reference belong to the discussion about the theories of meaning. The theories of meaning are concerned with the meanings of linguistic expressions and about how those meanings came into being. Theories of reference cover a subsection of this and are concerned with:

a) how (certain types of) linguistic expressions we use – proper names, natural kind terms, definite descriptions, etc. refer to (objects in) reality;

b) what is the connection, if any, between reference and meaning and c) reference and truth.

Traditionally, a *descriptivist* theory of reference dominated. Kripke (1980) has criticized the descriptivist views by pointing out several serious shortcomings, and proposed an alternative approach to reference, arguing for its validity by presenting several convincing thought experiments. His view was later dubbed as *causal-historical* theory of reference, and has dominated the philosophical world ever since. Recent studies in experimental philosophy raise doubts about the successfulness of Kripke's refutation of descriptive theory, and the potential successfulness of any theory of reference to apply universally. In the next chapters I will give an overview of the content and the issues with the two aforementioned theories of reference. I will also describe a third option which is trying to combine the first two, seemingly mutually exclusive theories.

2.1.   Descriptive theory of reference

An earlier theory of reference, argued for by Mill, considers the meaning of a proper name to be its bearer, so that proper names denote, but do not connote. The reason why Mill's

theory is not sufficient for a proper explanation of reference is its failure to make sense of a sentence, if this sentence would:

a) use a name, a bearer of which does not exist in the real world, e.g. "Spider-Man is brave";

b) deny the existence of something which does not exist in the real world, e.g. "Spider-Man does not exist";

c) use different names for the same bearer in a conflicting manner, e.g. "My friend believes Lewis Carroll is a great writer, but that Charles Lutwidge Dodgson is a lousy one".

Frege presents a theory which would solve these issues. According to this theory the reference of an expression consists of descriptions the speaker associates with that expression, and this description picks out the referent uniquely. According to Frege, several expressions can have the same reference, i.e. (actual) object corresponding to them, while having different senses (meanings). Also, an expression can have a sense (i.e. it can be understood) without having a reference (such as Spider-Man in the earlier examples). This all applies to expressions; in case of sentences, the reference is its truth value. Russell[1], continuing the descriptive tradition, claims that proper names are actually abbreviated definite descriptions, and the reference of an expression is definite, consisting of one or more descriptions combined.

Mallon *et al.* (2009, p. 334) describe the generic form of a simplified descriptive theory as follows:

> D1. Competent speakers associate a description with a term *t*. This description specifies a set of properties.

> D2. An object is the referent of *t* if and only if it uniquely or best satisfies the descriptions associated with it.

Let us list, as an example, a set of descriptions for a proper name *Tartu*: a) the second largest city in Estonia, b) the home for 98 393 citizens as of 1st of January 2010, c) the location for Estonia's oldest university, founded in 1632 by the Swedish king Gustavus Adolphus.

---

[1] Russell, Bertrand. 1905. "On Denoting"

Although the descriptive theory overcomes the problems present in Mill's theory of reference, there are serious issues with this theory itself. Kripke, one of the most notable critics of the descriptive theory, has tackled several of these in his (1980).

Firstly, if, as it follows from the descriptive theory, a description is similar to a proper name and therefore a *rigid designator*, which holds in all possible worlds, then the description would become *necessary*. So, if Tartu was indeed a home for 98 393 of its citizens on 1st of January 2010, it is necessarily so, and in all possible worlds. It is an easily imaginable scenario, though, that Tartu had exactly 100 000 citizens on 1st of January 2010 in some other possible world. The name can be the same in all possible worlds, but the descriptions differ. *Ergo*, descriptions do not equal names.

Secondly, if only a limited set of details is known about a name's reference, the description could fail to uniquely pick out the correct reference. If all what we have for a description is "the home for 98 393 citizens as of 1st of January 2010" then the reference would fail as there can be several cities in the world, exactly with the same number of citizens.

Thirdly, the description that is commonly used for a name – borrowing from Kripke e.g. when Einstein is referred to as the inventor of the atomic bomb – could be wrong.

Kripke gives his own example to illustrate some of these issues in a form of a thought experiment (Kripke 1980, pp. 83-84):

> Let's suppose someone says that Gödel is the man who proved the incompleteness of arithmetic, and this man is suitably well educated and is even able to give an independent account of the incompleteness theorem. [...]

> Let's take a simple case. In the case of Gödel that's practically the only thing many people have heard about him-that he discovered the incompleteness of arithmetic. Does it follow that whoever discovered the incompleteness of arithmetic is the referent of 'Gödel'? [...]

> Imagine the following blatantly fictional situation. [...] Suppose that Gödel was not in fact the author of this theorem. A man named 'Schmidt', whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the view in

question, then, when our ordinary man uses the name 'Gödel', he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description, 'the man who discovered the incompleteness of arithmetic'. Of course you might try changing it to 'the man who published the discovery of the incompleteness of arithmetic'. By changing the story a little further one can make even this formulation false. Anyway, most people might not even know whether the thing was published or got around by word of mouth. Let's stick to 'the man who discovered the incompleteness of arithmetic'. So, since the man who discovered the incompleteness of arithmetic is in fact Schmidt, we, when we talk about 'Gödel', are in fact always referring to Schmidt. But it seems to me that we are not. We simply are not.

Here, the descriptive theory would force us to have Schmidt as the reference and causal-historical theory allows us to use Gödel as the reference. As the descriptive theory seems to fail in this case Kripke proposes an alternative theory, a "better picture", which would solve the issue.

2.2. Causal-historical theory of reference

Causal-historical theory of reference explains the reference for proper names through fixing (in other words, initial baptism), where the first use of a name is followed by a sequence of causal links of the use of the same name. This kind of link between a referent and its description is *contingent*.

Here is how Mallon *et al.* (2009, p. 335) describe the generic form of a simplified causal-historical theory:

C1. A term *t* is introduced into a linguistic community for the purpose of referring to a particular thing (e.g. a person or a property). The term continues to refer to that thing as long as its uses are linked to the thing via an appropriate causal chain of successive users: every user of the term acquired it from another user, who acquired it in turn from someone else, and so on, back to the first user who introduced the term.

C2. Speakers may associate descriptions with terms. But after the term is introduced, the associated description does not play any role in the fixation of the referent. The referent may entirely fail to satisfy the description.

The proposed approach does not solve all issues with reference. The rules of how a name is baptised are flexible, so a name heard in a conversation can be meaningfully used by the person who overheard it, and the reference would still be correct, in the sense of following the chain of borrowing. The causal-historical theory in its primary form cannot fully account for changes in meaning that can historically occur (e.g. "Madagascar"[2]), and which lead to a change in the reference of the proper name. Proponents of causal-historical-type of theories of reference have extended the theory by allowing reference changes over time, as these are historically fixed by the new use of the proper name.

2.3.   Other options

Both the descriptive and the causal-historical theory of reference have their opponents and proponents and the jury is still out on which of these is correct, or if either of them can be used individually as a satisfactory theory of reference. Several alternative theories of reference combine the strengths of both of the discussed theories, thus overcoming the described issues. For instance, Evans provides a hybrid theory of reference, according to which the reference of a proper name is determined by the dominant causal source of the body of descriptive information the speaker associates with the name. This allows people to refer to objects they have no causal link to whatsoever. Devitt argues that the first part of the causal-historical reference basis – the baptism – is done descriptively, as the baptiser has to have some sort of description for the object to be baptised; and the second part – the borrowing – follows the historical chain of proper name usage.

The preference of a causal-historical theory of reference over the descriptive has been common among philosophers since Kripke's famous argumentation, though. Kripke used convincing thought experiments as a part of his argumentation. Machery *et al.* (2004) have

---

[2] "Gareth Evans has pointed out that similar cases of reference shifts arise where the shift is not from a real entity to a fictional one, but from one real entity to another of the same kind. According to Evans, 'Madagascar' was a native name for a part of Africa; Marco Polo, erroneously thinking that he was following native usage, applied the name to an island. (Evans uses the example to support the description theory; I, of course, do not.) Today the usage of the name as a name for an island has become so widespread that it surely overrides any historical connection with the native name [...] So real reference can shift to another real reference, fictional reference can shift to real, and real to fictional. In all these cases, a present intention to refer to a given entity (or to refer fictionally) overrides the original intention to preserve reference in the historical chain of transmission."(Kripke 1980, p. 163)

argued that these thought experiments are successful, at least partly, due to their intuitive correctness – which is an issue in case intuitions turn out not to be universally shared.

To test their hypotheses of variance in intuitions regarding the theory of reference, Machery *et al.* conducted experimental studies using probes, inspired by Kripke's thought experiments. These probes were meant to elicit intuitions from the test subjects specifically regarding the questions of reference. Based on the results of these studies, Machery *at al.* claimed that intuitions regarding reference are indeed different for different demographic groups, and as such, do not make good ground for any philosophical argument.

If we accept Machery *et al.*'s description of the analytic philosophy's argumentation methodology as an intuition-based armchair activity, and agree with the conclusions they have drawn from the findings of their experiments, a significant challenge is created for this "armchair philosophy".

Machery *et al.* are not alone on their quest. Several other experimental philosophers have conducted similar studies, although their goals and conclusions do not always exactly match those of Machery *et al.* The aforementioned challenge – the experimentalist challenge – is generally supported, though, and in the next chapter I set out to describe the significance of this challenge in detail.

3. The experimentalist challenge

Weinberg (2009) describes the *experimentalist's challenge* as a challenge to the armchair practices of philosophy. The research by experimental philosophers (or *experimentalists*, as Weinberg calls them) has shown the basis for judgements as deployed by philosophers – i.e. intuitions – to not be universal and armchair-accessible, but (potentially) sensitive to various factors usually considered irrelevant to finding philosophical truth, such as demographic differences, order effects, framing effects and environmental influences (p. 456). These factors are often covert even for the thinker, which makes them hard to trace.

Machery *et al.* (2004) is a prime example of the early experimental research, and is the motivation and basis for several following studies. Machery *et al.* conducted an experiment to test demographic differences in intuitions about reference, results of which supported their hypothesis of variation between the intuitions (about reference) of Westerners and East Asians. The details of this research will be discussed in Chapter 3.4.

The experiments conducted by experimental philosophers since the initial Machery *et al.*'s study are diverse, and while the motivation behind some of these is to find out which of the philosophical judgements are sensitive to the aforementioned variation and which are safe, other experiments are designed to show that error-prone intuition-based judgements should be excluded from acceptable philosophical practice altogether.

In the next chapters I will describe the overall framework and different branches of experimental philosophy to show the context for the Machery *et al.*'s study. I will then discuss the alleged importance and pervasiveness of intuitions in philosophy by contrasting Machery *et al.*'s views with some of their critics'. I will provide my detailed argument for the view that Kripke, although perhaps undeservedly singled out among the "armchair philosophers", has indeed considered intuitions as suitable evidence in philosophical

discussions, and that the experimental challenge is real and relevant. Finally, I will present my response to the experimentalist challenge.

3.1.   The framework of Experimental Philosophy

Experimental philosophers have applied various methods in their research, but there are some common traits to most experimental philosophers. Nadelhoffer and Nahmias (2007, p. fn 141) list two main criteria for one to be considered an experimental philosopher:

1) One runs controlled and systematic studies and uses the resultant data to shed light on philosophical problems;
2) One sometimes addresses the tension that exists between what philosophers say about intuition and human cognition, on the one hand, and what researchers are discovering about these things, on the other hand.

So we can say that a true experimental philosopher conducts studies and analyzes the data in relation to intuitions and cognition. Recent trends in experimental philosophy have shown a slight deviation from these criteria. As experiments have become manifold, the amount of raw data available for analysis has increased significantly. The quality of research is at times disputable and although the experiments themselves are occasionally severely criticized   from   the   methodological   point   of   view

[3], resulting in rendition of some of the gathered data invalid for any relevant (philosophical) analysis, there are "theoretical experimental philosophers", who are building their arguments on already available experimental data, and (sometimes) arriving at different conclusions from the ones of the "hands-on experimental philosophers". Based on this, I would add the following third criterion to the experimental philosopher's specification:

3) In the absence of their own unique studies, one utilizes existing data by scrutinizing the methods used for both the gathering and previous analysis of this data.

---

[3] See Chapter 5 for detailed overview of the methodological criticism

The philosophers satisfying these three criteria differ in their motivation for experimental studies. I have divided the existing research and positions in experimental philosophy into two motivational categories – *explorative* and *sceptical*.

3.2.   Explorative Experimental Philosophy (EEP)

According to EEP[4], intuitions can be a trustworthy source of evidence for forming justified beliefs. Borrowing from Nadelhoffer & Nahmias (2007), I will divide EEP into two sub-categories.

(A) Experimental Analysis explores, in a controlled and systematic manner, what intuitions ordinary people tend to express and examine their relevance to philosophical debates. This can verify or refute philosophers' claims that their position aligns with common sense. Practitioners of Experimental Analysis believe that all intuitions, including those of the folk, are relevant to establish the content for concepts.

(B) Experimental Descriptivism explores the nature and sources of philosophers' intuitions, the role that they should play in philosophy and how to best explore them. Practitioners of Experimental Descriptivism think that intuitions should not be dismissed in the philosophical practice, but do not consider folk intuitions relevant.

For example, Alexander (forthcoming) thinks that we do not yet know what counts as intuitions. The content of the concept is not fixed yet and experimental philosophers, when discussing intuitions, are operating with "thin concepts", as he puts it. By using these thin concepts to explore how similar concepts have been historically used by philosophers, the borders become clearer and the content for the concepts more fixed.

3.3.   Sceptical Experimental Philosophy (SEP)

According to SEP[5], armchair reflection and informal dialogue are not reliable sources of evidence for (philosophically relevant) claims about folk concepts. Practitioners of SEP

---

[4] Same as *Optimistic Experimentalism* (Kauppinen 2007); *Positive Program* (Alexander, Mallon & Weinberg forthcoming)

conduct surveys with folk to show the inconsistency of intuitions and their sensitivity to various factors – culture, socio-economical status, gender, etc. It is difficult to assert, objectively, why anyone's intuitions should be held in higher regard than those of others, if the diversity occurs. As the intuitions are inconsistent and unreliable, it is impermissible to use them to theorize about philosophical topics. This position does challenge the Explorative program, as Alexander (forthcoming) also notes. It can be considered too radical and unnecessary, and alternative "healthy and already sufficient" views can be proposed (Liao 2008). On the other hand, if the sceptics are successful in their enterprise and experiments' data supports their claims, it leaves no room for intuitions as evidence in philosophy.

3.4.   The Machery study

Machery *et al.*, representatives of Sceptical Experimental Philosophy, believe intuitions to be the main basis for many philosophical arguments, and certainly so for the arguments about reference. According to them, "philosophers agree that theories of reference for names have to be consistent with our *intuitions* regarding who or what the names refer to" (Machery et al. 2004, p. B2). They believe that Kripke's argumentation, refuting the descriptivist theory of reference, is successful precisely because it elicits intuitions inconsistent with the descriptive theory of reference. Referring to experimental research in cultural psychology, e.g. (Nisbett 2003) on the topic of cognitions, and in philosophy, e.g. (Weinberg, Nichols & Stich 2001) on the topic of epistemic intuitions, Machery *et al.* argue that although it is the case for most philosophers in regard to reference, it is unwarranted to assume (a cultural) universality of cognitions and intuitions, as for both significant differences have been found between East Asians (EAs) and Westerners (Ws)[6]. Machery *et al.* set forth the suspicion that similar variation can exist in intuitions about reference of proper names, specifically that EAs would be more likely to respond according to the descriptive theory of reference compared to Ws, as the aforementioned experiments in psychology have shown Ws' inclination over that of EAs' to base their reasoning and judgements on causality. To verify this hypothesis

---

[5] Same as *Pessimistic Experimentalism* (Kauppinen 2007); *Experimental Restrictionism*(Alexander, Mallon & Weinberg forthcoming) and (Nadelhoffer & Nahmias 2007); *Radical Experimentalism* (Liao 2008)
[6] East Asians participating in experimental philosophers' research are usually Chinese, Japanese, Korean; participating Westerners are usually Americans.

empirically using an experiment of their own, they designed several probes, based on Kripke's thought experiments, and conducted the experiment with students from US (representing Ws) and Hong Kong (representing EAs).

After eliminating the demographically unsuitable participants from the experiment data, the responses of 31 Western undergraduates from Rutgers University and 40 Chinese undergraduates from Hong Kong University were analyzed from the initial total of 82 participants. The participants had been presented with four probes, two of which were modelled on Kripke's Gödel case and two on Kripke's Jonah case. Both types of probes had one version with English names and the other with Chinese names. In the quoted probes below, the answer in line with the descriptive theory of reference is (unbeknownst to the participants during the experiment) marked as (A) and the answer in line with causal-historical theory of reference is marked (B).

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

> Jonah case (Chinese-name version):
> Lau Mei Ling is a high-school student in the Chinese city of Guangzhou. Like everyone who goes to high-school in Guangzhou, Mei Ling believes that Chan Wai Man was a Guangdong nobleman who had to take refuge in the wild mountains around Guangzhou in the eleventh century A.D.,

> because Chan Wai Man was in love with the daughter of the ruthless Government Minister Lee, and the Minister did not approve. Everyone in Lau Mei Ling's high-school believes that Chan Wai Man had to live as a thief in the mountains around Guangzhou, and that he would often steal from the rich allies of the Minister Lee and distribute their goods to the poor peasants. Now suppose that none of this is true. No Guangdong nobleman ever lived in the mountains around Guangzhou, stealing from the wealthy people to help the peasants. The real facts are the following. In one of the monasteries around Guangzhou, there was a helpful monk called "Leung Yiu Pang". Leung Yiu Pang was always ready to help the peasants around his monastery, providing food in the winter, giving medicine to the sick and helping the children. Because he was so kind, he quickly became the main character of many stories. These stories were passed on from one generation of peasants to the next. Over the years, the story changed slowly as the peasants would forget some elements of the story and add other elements. In one version, Leung Yiu Pang was described as a rebel fighting Minister Lee. Progressively the story came to describe the admirable deeds of a generous thief. By the late fourteenth century, the story was about a generous nobleman who was forced to live as a thief because of his love for the Minister's daughter. At length, not a single true fact remained in the story. Meanwhile, the name "Leung Yiu Pang" was slowly altered: it was successively replaced by "Cheung Wai Pang" in the twelfth century, "Chung Wai Man" in the thirteenth, and finally by "Chan Wai Man". The story about the adventurous life of Chan Wai Man was written down in the fifteenth century by a scrupulous historian, from whom all our beliefs are derived. Of course, Mei Ling, her classmates and her parents know nothing about these real events. Mei Ling believes a story about a generous thief who was fighting against a mean minister.

> When Mei Ling says "Chan Wai Man stole from the rich and gave to the poor", is she actually talking about the generous monk, Leung Yiu Pang, who is the original source of the legend about Chan Wai Man, or is she talking about a fictional person, someone who does not really exist?

> (A) She is talking about the generous monk, Leung Yiu Pang.
> (B) She is talking about a fictional person who does not really exist.

The findings are harmonious with the results of a pilot study the authors had conducted with 19 Western and 32 Chinese participants earlier. The statistical analysis of the experiment's results showed some cultural variation in the responses for the Gödel cases, as Machery *et al.* had predicted. The Westerners chose causal-historical (Kripkean) responses over descriptive, while the Chinese students leaned towards descriptivist responses. Contrary to their speculations, no significant variation was found in Jonah cases. Seemingly

because the results for the Gödel cases are in line with authors' predictions, they do not discuss the latter (non-)finding regarding the Jonah cases in detail, but do provide some speculations why it is so – they suspect it is so either due to difficulties arising from the excessive length of the probe's text, or due to pragmatic reasons where the Chinese participants chose the causal-historical option over the descriptive where the speaker's term might have otherwise failed to refer at all. In their main line of argumentation, the authors choose to ignore this difference between two types of probes. As it has been pointed out by several critics, Machery *et al.* celebrate their findings too early, as methodological analysis of both types of probes would have been in order once potential issues were discovered with the Jonah cases. Machery *et al.* fail to run this analysis, though.

While acknowledging the inconclusiveness of their experiment and the need for future research on the subject, Machery *et al.* interpret the findings to show that "it is wrong for philosophers to assume a priori the universality of their own semantic intuitions [regarding the referents of terms]" (p. B8). This interpretation depends, of course, on a premise that philosophers have traditionally indeed been assuming the universality of intuitions in their argumentation.

If this proves not to be the case and the universality of all intuitions is not required, a distinction has to be defended why only some intuitions are treated as relevant in philosophical discourse. The experiments studied folk intuitions, and Machery *et al.* speculate that philosophers could potentially draw the line by claiming that only reflective philosophically-informed intuitions can help with finding the correct theory of reference, not those of ordinary people. The authors do not find this convincing as they cannot find any solid argument to prefer Western academic philosophers' intuitions to those from any other cultural or linguistic group, especially as these philosophers' intuitions could be contaminated by not being independent, but reinforced by the traditions of philosophical training. Another interpretation of the role of intuitions in philosophy, according to which they would help to "develop an empirically adequate account of the implicit theory that underlies ordinary uses of names" (p. B9) would still require the (cultural) variation to be taken into account as reference would, based on the experiment's results, work differently for different (cultural) groups.

In addition to cross-cultural variation, the authors also discovered intra-cultural variation, as the deviation in answers within both test groups was also significant (p. B8). Machery *et al.* speculate that these findings can also mean variation between smaller groups than cultural, or perhaps even on individual level. In their (2004), the authors did not set out to discover and analyze all possible variations in intuitions and their explanation, but to show that some variation does exist. The significance of their findings is that they cast a shadow over the (allegedly) intuition-laden methods and practices of analytic philosophers, which have now, as a response to experimental philosophy, been defended perhaps even more rigorously[7] than ever before.

The choice of intuitions about the theories of reference for their experiment was not arbitrary, as they discuss with further analysis on intra-cultural variation, in (Mallon et al. 2009). They claim that philosophers often make use of a specific intuition-laden method when contemplating the correctness of a theory of reference – namely, *the method of cases* (p. 338):

> *The method of cases:* The correct theory of reference for a class of terms T is the theory which is best supported by the intuitions competent users of T have about the reference of members of T across actual and possible cases.

It is this method, according to the authors, which leads Kripke to refute the descriptivist theory of reference, and other philosophers, including now-former descriptivists, to agree with him. The intuitions of the latter, as the authors note, changed from descriptivist to causal-historical once suitable counterexamples were constructed by Kripke, which shows the inconsistency of one's intuitions over time, even in case of philosophers. The distinction between "correct" and "incorrect" theory of reference is not clear-cut, as a philosophically accepted (universal) variation of intuitions in regard of the classes of terms in question can be argued for, with some classes eliciting causal-historical, and some descriptive intuitions. The findings in (Machery et al. 2004), on the other hand, describe cross- and intra-cultural variation of intuitions regarding the same classes of terms – something that has not been seriously considered by philosophers so far.

---

[7] e.g.(Kauppinen 2007)

This extensive variation requires a response, and the authors propose three options. According to the first option, there can be no arguments from reference[8], as we do not and cannot have a correct substantive theory of reference considering the diversity in intuitions about reference (p. 342). According to the second option, it can be argued that the role of intuitions (in determining the correct theory of reference) is less significant than the method of cases proposes. The authors do not consider this a relevant option as arguments from reference begin with independently motivated theory of reference – something that is appealing to intuitions – and arguments motivated in any other way are not arguments from reference in their use of the term, therefore irrelevant to the current discussion (p. 343).

The third option describes a potential theory of reference, *referential pluralism*, according to which the variation in intuitions about reference is caused by the actual differences in reference for different groups. It can lead to apparent contradictions, as a proposition about a concept, thought to be universal, can be true and false at the same time for members of different *intuition groups* – groups of persons sharing intuitions about a set of cases. This, as they note, is not a good option, as the number of intuition groups is unpredictable due to discovered cross- and intra-cultural variation, and this leads to *ubiquity of variation*. As a possible solution, the propositions could be relativized, in which case the content of the same utterance can be different for members from different intuition groups. This in turn would require the resulting philosophical arguments to be relativized as well. Genuine disagreement between philosophers, and any basis for arguments, can exist only when they belong to the same intuition group. This relativist view will lead to absurd conclusions as genuine disagreement is difficult to prove: a lot of cases have to be considered for each and every speaker to determine the intuition group the speaker belongs to, and one's explicit views about their group membership cannot be trusted due to being inconsistent, as shown above regarding the change of the acceptable theory of reference. This is the experimental challenge, leaving philosophers, according to Mallon *et al.*, at crossroads. They need to either abandon the arguments from reference, as the only source of getting information about the reference are intuitions and these cannot be trusted, or to hold their ground and deny the

---

[8] "[...] arguments that derive philosophically significant conclusions from the assumption of one or another theory of reference" (Mallon et al. 2009)

variation in intuitions about reference, which would eliminate the problems with the widely used method of cases.

The aforementioned second option, according to which the role of intuitions in philosophical argumentation is less significant than stated by experimental philosophers, is not universally dismissed, and has appealed to several philosophers as a way to explain away the philosophical significance of discovered variation in intuitions. Max Deutsch has recently published two well-structured articles on this subject and I will discuss his argumentation in detail in the following chapter.

3.5.   The role of intuitions in philosophy: analyzing Kripke

It has been argued that the role of intuitions in analytic philosophy, and specifically in Kripke's argumentation against the descriptive theory of reference, is overstated, or misunderstood by the experimental philosophers. I would like to present this line of criticism, and argue that the evidence provided by the critics for their interpretation of Kripke is insufficient to battle that of experimental philosophers.

Deutsch (2009) and (2010) criticises experimental philosophers' position about the role of intuitions in philosophy, and their interpretation of Kripke's arguments harshly. He argues that experimental philosophers have wrongly considered intuitions as something on which philosophical argument traditionally depends, and have mistakenly attributed this position to Kripke as their chosen example of this approach. Here are Deutsch's (2009) arguments:

a) Counterexamples, also used by Kripke, are widely used in philosophy and other sciences as a common method to refute conclusions based on generalizations;

b) Kripke's counterexamples do not rely on intuition: "A counterexample is presented with no explicit or implicit appeal to its intuitiveness";

c) The discussion about the intuitiveness of the counterexamples in Kripke's case is irrelevant: "Whether these counterexamples are intuitive for anyone is a separate, and purely psychological, matter";

    d) Kripke succeeds in refuting descriptivism by using only the method of counterexamples, not by relying on intuitions: "Kripke refuted descriptivism, if he did, by presenting counterexamples, full stop".

Following from this, if the only thing that matters in the argumentation about the theory of reference, in Kripke's case, is the successfulness of the counterexamples and philosophers' intuitions are irrelevant to the argumentation, then any empirical data about folk intuitions, gathered by conducting experiments, is also irrelevant to the argumentation. If experimental philosophers, including Machery *et al.* (2004) and Mallon *et al.* (2009), have made a mistake in choosing Kripke as an example of a philosopher whose belief in his own intuitions might cloud his judgement, then the their case against armchair philosophy and the relevance of (folk) intuitions regarding the theory of reference is seriously misguided. Before I discuss the method of determining the successfulness of counterexamples, I want to show that Kripke's reliance on intuitions in his argumentation is not as clearly nonexistent as Deutsch argues, and that experimental philosophers have not made a mistake when choosing Kripke and his influential thought experiments as a target for their inquiry.

By analyzing Kripke's writings, one can see that Kripke held intuitions in high regard, and the following quote is an example of this (Kripke 1980, p. 42):

> Of course, some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking.

When describing his overall attitude towards intuitions, Kripke is using the phrase "[most] conclusive evidence" to describe the status of intuitions. He is giving a strong justificatory role to intuitions. The word "evidence" has been used widely in discussions within experimental philosophy and I argue it has been used somewhat carelessly. To show the reason for this concern, a formal distinction between *intuitions*, *evidence*[9] and *proof*[10] is in order. Philosophical intuitions have mainly been used in two distinct ways. The first practice

---

[9] Merriam-Webster dictionary: „evidence: something that furnishes proof"
[10] Merriam-Webster dictionary: „proof: the cogency of evidence that compels acceptance by the mind of a truth or a fact; something that induces certainty or establishes validity"

gives intuitions the status of proof, i.e. for something to be intuitive means it to be true. This is the way Kripke has used intuitions, although he uses the word "evidence", not "proof". One can speculate why Kripke has chosen to burden the word "evidence" with so much strength. As mathematics is also discussed by him in the same context, and "proof" in that specific context has a very clear non-arguability and *a priori* content, he might have wanted to keep intuitions and *a priori* separate. The way he uses them, though, gives them a similar *a priori* status of mathematical intuitions, being the final non-empirical proof. The second practice, argued for by practitioners of Explorative Experimental Philosophy[11], states that intuitions can be used as exhibits in the philosophical trial for or against any arguments, and in this way intuitions can be at most evidence, but not (final) proof.

With some variation, Deutsch agrees with the second practice. He does not deny that intuitions have been used in philosophical argumentations in general, but argues that their role has been misunderstood. He distinguishes between the *justificatory source* and the *causal source* for judgements. Intuitions, according to him, cannot be and have not been used as a justificatory source (which he describes as evidence), but only causal source. Deutsch's concept of evidence differs from the evidence and proof distinction described above. It is clear that there is confusion between concepts. Sceptical Experimental Philosophy argues against using intuitions even as evidence in the "can be exhibits" sense, as intuitions are thoroughly unreliable. Explorative Experimental Philosophy on the other hand allows intuitions to be evidence (i.e. exhibits), but does not grant them the status of proof – they are just one way of adding knowledge. By treating experimental philosophy as a unified enterprise, which argues against using intuitions as evidence in Deutsch's sense, he equates experimental philosophy with Sceptical Experimental Philosophy, something with which a large number of experimental philosophers (e.g. practitioners of Explorative Experimental Philosophy) would not agree.

Did Kripke, in addition to this general attitude towards intuitions, consider them to play a role in the much-quoted Gödel thought experiment, or have experimental philosophers misread it? Deutsch thinks intuitions were not referred to, and describes Kripke's wording

---

[11] And supported, among others, by Deutsch, Devitt and Liao.

of the argument as something being said "straight out, and emphatically" (Deutsch 2009, p. 451). Here is the relevant section from Kripke (1980, p. 84):

> On the view in question, then, when our ordinary man uses the name 'Gödel', he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description, 'the man who discovered the incompleteness of arithmetic'. [...] So, since the man who discovered the incompleteness of arithmetic is in fact Schmidt, we, when we talk about 'Gödel', are in fact always referring to Schmidt. But it seems to me that we are not. We simply are not.

I disagree with Deutsch in describing this as something being said "straight out". The keyword I'd like to draw attention to in this passage is "seems". Kripke does not specify what the proof is for us not referring to Schmidt in this example. What he says is that for him *it seems* that we are not referring to Schmidt. Deutsch manages to miss (or avoid) this part of Kripke's sentence in both of his articles.

In English, as used by competent speakers as well as scientists when talking about observations or conclusions, the verb "seem" denotes a feeling or belief about something which is not supported by facts. This interpretation is also argued for by Grice, when he describes the use of the verb "seem" in cognition as "noncommittal" and "subjective" (Grice 1991 (1989), p. 140):

> Standardly the generated communication will be an informal one to the effect that the use of the word "seem" is well chosen in relation to that of some identifiable contrasting expression to which it is preferred. [...]
> First, distinguishing three cases of contrasting terms:
> (1) Low subjective contrasters
> (2) High subjective contrasters
> (3) Objective contrasters
> Examples of (1) could be "It seems to X that p," "X thinks that p," or "It looks to X as if p." Examples of (2) could be "X knows that p," "X sees that p," possibly "It is clear to X that p," or "It is apparent to X that p." In (3), whatever condition is expressed by "p," the objective contraster would be the sentence saying what p is.

One can argue that Kripke did not *intend* to convey this Gricean meaning with his phrasing and it was a mishap, but his position regarding intuitions, as described above, suggests otherwise.

By using intuitions as the "most conclusive evidence", traditional philosophy is open to attack by the sceptics, who demand further (empirical) proof, and now have some of the experimental philosophers in their ranks.

Even if Kripke was using intuitions in his argumentation, his actual accomplishment in showing the defectiveness of descriptive theory, according to Deutsch, did not rely on intuitions, but on successful counterexamples. We are now left with a question about how to *evaluate* the successfulness of counterexamples. Deutsch extends his argumentation by describing how he thinks this can be done.

For a counterexample to be successful, it has to be *genuine*. There is not necessarily only one basis for this evaluation. There can be various grounds on which to judge a counterexample as genuine, but it is certain intuitions are not one of those.

Deutsch defines two types of grounds – the K-Grounds (the reasons for granting Kripke's Gödel case the status of being genuine) and G-grounds (the reasons for granting the Gettier cases against knowledge as justified true belief the status of being genuine). Following is a description of the K-Grounds (Deutsch 2010):

(a) Kripke points out that the imaginary Gödel-case has real-life analogues. All that many of us 'know' about Peano is that he was the discoverer of certain axioms concerning the natural numbers. But it turns out that Dedekind discovered those axioms. If descriptivism is true, many of us have been referring all along to Dedekind with our uses of 'Peano'. But we have not been referring to Dedekind with those uses. We have been referring instead to Peano, misattributing to him the discovery of the axioms. This is not simply a further putative counterexample; it strengthens the claim that the Gödel-case is a counterexample by showing us that the way in which we ought to judge, with respect to the imaginary Gödel-case, should line up with the way in which we correctly judge about the real-life Peano case. (See Kripke 1980, 84-85.)

(b) Kripke argues that the view that 'Gödel' refers to Schmidt—the prediction made by descriptivism concerning the Gödel-case—suggests a more general view to the effect that one can never be mistaken in uttering a sentence of the form 'N is the F', when 'the F' denotes, and is a definite description one associates with 'N', a proper name. But one can be mistaken in uttering 'Peano is the discoverer of the axioms', even if one associates 'the discoverer of the axioms' with 'Peano'. The falsity of this general view is evidence that Kripke is right in claiming

that 'Gödel' refers to Gödel, not Schmidt, in the Gödel-case. (See Kripke 1980, 85n, 87.)

(c) Kripke argues for an alternative account of the way in which 'Gödel' refers (the causal-historical account) which explains, Kripke thinks, why 'Gödel' refers to Gödel in the Gödel-case. The existence of a satisfying general theory of reference that predicts that 'Gödel' refers to Gödel in the Gödel-case counts in favor of the view that 'Gödel' refers to Gödel in the case. (See Kripke 1980, 91–93.)

I will show why I do not find these grounds convincing. In ground (a) Kripke claims that we have not been referring to Dedekind with our uses of "Peano", but to Peano by a misattributing a certain description to him. The first problem is that he does not explain why he thinks so. How does he know to whom people have been referring? Is this *a priori* knowledge, or what empirical data did he use to make this conclusion? I can agree with him intuitively, because it seems to make sense, but I cannot provide any reasonable argument to why it is so, and neither can Deutsch. Secondly, I see that an alternative explanation which supports the descriptive theory of reference can be envisaged. This explanation would place the mistake to the act of referring, instead of the act of attributing, and in this case we have indeed been referring to Dedekind by using "Peano", contrary to what Kripke is claiming. How is this possible? If the majority of speakers are making an error in their communication, on regular basis, the error "disappears". Let's say that at first, nobody knows it was Dedekind who discovered the axioms. Everybody believes it was Peano. When people (mistakenly) used "Peano" for the "discoverer of axioms", it worked well on the conversational level because it was a shared mistake. Not a mistake from people's point of view at that time, but a mistake in the reference, i.e. this was not true. If there would have been a very small group of people, who knew the truth and were using "Dedekind" in the sense of "discoverer of axioms", they would have been correct and referring correctly, but mistaken conversationally when speaking to the non-knowers. Now, when it becomes common knowledge that Dedekind was indeed the discoverer of the axioms, not Peano, everybody using "Peano" in the sense of "discoverer of axioms" would be mistaken both in reference and conversationally. What this explanation argues for is that how people actually refer can conflict with what is the correct reference. This explanation seems as intuitively correct for me as the one provided by Kripke. It seems to me this alternative explanation

also takes care of ground (b), as the mistakes can be different – from the point of truth value, and conversationally. Similar situations can occur and have occurred in science relating to scientific concepts, such as "compound", where the original reference is later discovered to have been false (Fine 1975).

It is difficult to determine, what exactly counts as a justification for ground (c). Deutsch acknowledges that the grounds can require more proof (Deutsch 2010):

> It is true that the K-Grounds and G-Grounds appeal to various principles and propositions for which we may demand further grounds. [...] However, it is implausible to suppose that the last link of every justificatory chain must be a premise asserting the intuitiveness of some proposition or other. [...] More often than not, justifications come to an end with premises that assert something other than that some proposition is intuitive for someone or some group of people. [...] I can imagine experimental philosophers complaining that, although there need be no rock bottom appeal to intuitions in arguments for philosophical counterexamples, the fact is that there often is such an appeal. I have already conceded that there is a good deal of misleading talk about intuitions and their evidential role in philosophy.

He does not specify what "something other than intuitive" is, but claims it is "implausible" to think that it would be intuitions – at least partly because he claims to have shown that intuitions do not have the evidentiary role in philosophy that is assumed by experimental philosophers. We can agree that there are cases where no intuitions are involved and factual evidence – such as maps or encyclopaedias or similar – can be provided. For cases lacking the factual evidence, the question remains – what could the further grounds be? In other sciences, where counterexamples are also widely used, as Deutsch has previously noted, the justificatory chain can end with an empirical fact. Interestingly, empirical facts able to support the K-Grounds would have to be acquired by experimental philosophers – all methodological criticism towards their existing studies taken into account, of course. Could these improved studies ask about intuitions and still be relevant to the discourse?

I believe I have shown that Kripke considered intuitions more important than Deutsch argues, and without any other viable option presented so far, the "rock bottom" of the K-Grounds could also turn out to be an appeal to intuitions. If this is a problem for traditional philosophy or not, depends on if the findings by experimental philosophers of cultural

variations in intuitions can withhold the excessive methodological criticism targeted at their experiments, and if folk intuitions can be considered relevant to the philosophical discourse in the first place.

3.6.   Should we keep using intuitions in philosophy?

Before I discuss the relevance of folk intuitions to philosophical discourse, I want to show why we should keep intuitions usable as evidence in philosophy.

As I have previously noted, Alexander and Liao have both expressed their worries that Sceptical Experimental Philosophy, abolishing the use of intuitions from all philosophy, might be a stretch too much as it creates (only) destructive, rather than constructive, arguments. It has been argued that intuitions could still be used as evidence in philosophy, given that some conditions are met – for instance not considering philosopher's intuitions special in any way (Williamson 2004) – but as special have they been treated, when all other arguments have been exhausted and something to provide higher proof was needed in traditional armchair philosophy.

The reason why SEP can be so successful against this traditional practice is the difficulty for that practice to show that the intuitions philosophers have in the philosophical discourse are in any way different from folk intuitions; that they have some special psychological status and that they create special judgements. The philosophical thought experiments, for which judgements based on intuitions are elicited, do not differ considerably from actually occurring situations, and do not force our modal thinking into extremes. Real life examples can be presented for most thought experiments – as was it, for instance, with Kripke's Gödel and Peano cases. In reality, as Williamson argues, the philosophical intuitions have been used as a strange exception of intuitions in general, as they are not evidence about the truth content of a proposition P, but their appearance to one in the form "P is true" is in itself evidence for "P is true". It "[...] intellectually appears to one that P, not the fact that P itself" (2004, p. 119). Devitt believes this special status is unwarranted (2008, p. 8):

> To say that intuitions, whether those of philosophers or of the folk, are evidence is not to say that they are the only, or even the primary, evidence.

> Indeed, we can look for more direct, less theory-laden, evidence by studying what the intuitions are about, the reality itself; in particular, we can study the actual referential relation between names and their bearers.

Even if we agree that the reality itself is a much more reliable source for information than intuitions, can we still keep intuitions as a part of the evidence for a proposition?

To keep using intuitions as one of the sources of evidence for a proposition, at first we have to acknowledge that intuitions are of various strengths, argues Williamson. We have to give up the "(I intuit the proposition to be) true/false" approach to intuitions and replace it with "(I intuit the proposition to be) more likely to be true/false" and rate the intuitions' evidential status on a scale. To achieve the objectivity of the rating is not an easy task, as several factors, including our background beliefs and "wishful thinking"[12], can influence our own perception of the intuition's strength.

Secondly, we have to acknowledge that sceptics could keep asking for a justification of intuitive evidence *ad nauseam*. Alexander (forthcoming) distinguishes two types of sceptics – the *unpersuaded* and the *unpersuadable*. While the former are willing to accept certain kinds of reasons as "good enough reasons", thus enabling the use of intuitions, the latter are unwilling to accept that any such reasons exists. It would also be "good enough" for philosophy to provide suitable reasons for the unpersuaded critics and not get caught in the web of the unpersuadable sceptic's demands for proof.

The question if we have to give up the "dialectical theory of evidence" according to which philosophy should aim at rational persuasion with the help of evidence, or if we can keep it, depends on what we consider rational persuasion. In case we do not seek to satisfy the unpersuadable sceptics, as Alexander also recommends, we can continue using the dialectical theory. But, we have to give up treating intuitions as the source of (final) truth and solid knowledge. They help us to advance, but do not provide us with ready-made solutions or necessarily always something which can be called "facts about the world". Such fallibility of philosophy to provide ultimate and non-dismissible truth is not unique, though (Williamson 2004, p. 151):

---

[12] See Chapter 5 for more

> In philosophy as in empirical science, our evidence does not consist of facts with respect to which we are infallible. Ordinary knowledge is enough. We have no general guarantee against the possibility that we did not know something that we thought we knew. In philosophy, the evidence is even more contested than in empirical science.

If we give up intuitions as evidence, we would be giving up a significant method of acquiring new knowledge. What are the options for those who see this "soft" approach as problematic and are not satisfied with inconclusive results and potentially false knowledge?

Alexander (forthcoming) proposes that we could adopt a rather strict position of natural scepticism, according to which intuitions should not be appealed to as evidence in philosophy, because we just do not know for sure what intuitions are – we lack a naturalistically acceptable account of the cognitive faculty or faculties of intuition. But, as this leads to aforementioned loss of a knowledge acquisition method, and eventually the significance of (a large part of) experimental philosophy, this is not a sought for solution.

A lot of issues can potentially be avoided by claiming that intuitions are a source of *a priori* knowledge. In this case they could assume the role of proof, something that "establishes validity". If we set our objective to satisfy the naturalistically-inclined critics, this approach does not help much, as Devitt (forthcoming), shows. In his opinion *a priori* as a concept is not robust enough to be preferred to any other explanations. Although there are areas, e.g. mathematics, where knowledge seems not to be empirical – which means it could be considered to be *a priori* – it could much better be explained naturalistically by applying the holistic theory of justification. This option "undermines the motivation for the *a priori*" as something that is difficult to defend. Philosophers' intuitions, as this naturalistic theory describes, are not about the concepts, but about *kinds* – and these intuitions are empirical, based on a lifetime of experience with kinds. They do have a role in philosophy, but this role is not to provide *a priori* knowledge. Even logical truths, which are supposed to be *a priori*, are not necessarily so. The reason for this doubt is the lack of a satisfactory non-empirical account for the justification of logical truth, something that is required in the naturalistic theory. To be clear, Devitt does not argue that these truths are not, under any circumstances, *a priori* – he just prefers empirical explanation, which is more robust and has less (potentially lethal) counterarguments. The *a priori* is, as he says, still a mystery.

The overall role and importance of intuitions in philosophy might indeed be overstated by experimental philosophers. Machery *et al.* and Mallon *et al.* are especially responsible for this, but as I have shown in Chapter 3.5, intuitions have been used widely. If giving up intuitions as evidence in philosophical arguments leads us to a loss of a large part of philosophy, and considering intuitions as *a priori* knowledge does not work well because we still do not know what the *a priori* is, then perhaps we should not try to avoid them at all costs. What we need to do is acknowledge their fallibility and not require them to provide ultimate truth. Intuitions can help us gather evidence for a proposition, but the proof for the proposition has to be somewhere else – in the reality, as Devitt, Williamson and Alexander also agreed. This, as it has been argued, is not sufficient to answer the experimentalist challenge.

Weinberg (2007) lists four sources of hope – (a) external corroboration, (b) internal coherence, (c) detectability of margins and (d) theoretical illumination. These sources can provide error-detection and error-correction required for a method to be considered trustworthy. He is not convinced these can be applied to using intuitions as evidence, which is why intuitions are "hopeless" and should be dismissed from philosophical discussions: "Any putative source of evidence that is hopeless ought not to be trusted" (p. 327). His claims are only strengthened by Machery *et al.'s* studies which show significant variation in intuitions. In addition, he claims that specific person's intuitions can change over time, making the whole apparatus of intuitions highly unreliable.

Several answers can be given to this concern. Firstly, we can take a position that some concepts, such as reference, are constitutive, and their content is actually determined by the intuitions competent speakers have about them. I would not rule this out, but I believe it is extremely difficult to find the basis for distinction between constitutive and non-constitutive concepts, which is why I wish not to pursue this line of thought any further. Second answer is to question the studies on which Weinberg is building his arguments. The analysis by Machery *et al.* of their experimental data has indeed concluded that significant variation in intuitions exists, and it could possibly mean there is as many different intuitions about any topic as there are intuiters. The methodological criticism towards the studies, as will presented in chapter 5, creates doubt that any major conclusions could or should be drawn

based on that data. If the level of variance is actually significantly smaller than *Machery et al.* claim, or even non-existent, the intuitions are more uniform – and this leaves room for "traditional" shared intuitions, which can be discussed and checked against each other, until the "most correct" answer is found.

The third answer to the experimentalist challenge, which is also acknowledged by Weinberg (p. 337), is that recent studies in experimental philosophy[13] are studying brain activity and are more real-science-like, more what Weinberg prefers. At this point it is a minor set within the whole tome of experimental research, but it is growing.

The existence of proven (more) reliable methodology for studying intuitions – brain scans and Machery-type experiments with improved methodology – gives us "Weinberg-hope". There is no reason to give up on intuitions now, when we have just spotted the beacon.

---

[13] See e.g. Greene, J., and Haidt, J. 2002. How (and where) does moral judgment work? Trends in Cognitive Science 6 (12): 517–23. and Greene, J., L. Nystromm, A. Engell, J. Darley, and J. Cohen. 2004. The neural bases of cognitive conflict and control in moral judgment. Neuron 44: 389–400.

4.   The relevance of folk intuitions

From the previously presented information we can conclude that intuitions have been used in philosophy, sometimes to a large extent and by giving them a relatively high status. Philosophers' intuitions have helped them to present and prove their arguments. It is generally agreed that a prerequisite for identifying, being able to talk about and have intuitions about something is to have a concept of it[14]. In addition to using reference in everyday lives, both philosophers and folk, as competent speakers, can have a conceptual understanding of reference and therefore also have intuitions about it. If philosophers' intuitions about a concept differ, the reasons can be found out by clarifying the concepts in question to make sure the disagreement is indeed about the same concept, and then analyzing and discussing the matter between philosophers. They are educated folk – educated in philosophy, therefore experts in this field.

One of the main questions in the "intuitions do matter" view asks if philosophers' intuitions, granted that philosophers are experts in their field, are the only intuitions that matter for determining the correct theory of reference, or do the folk intuitions also hold (some) relevance to this discussion.

A common response[15] to the importance of the expert status is affirmative – philosopher's intuitions do indeed provide a higher value input for theories of reference. According to some philosophers, folk intuitions do not matter exactly because folk are not experts. The grounds for one to be counted an expert in the discussion about a concept do not come from just being a competent speaker of a language, as Deutsch for instance argues. He is very sceptical that non-expert folk intuitions, even those of undergraduates, about reference are relevant to understanding the concept of reference (2009, p. 453 fn 6):

---

[14] E.g. (Devitt 2006), (Kauppinen 2007)

[15] E.g. (Deutsch 2009), (Deutsch 2010), (Devitt 2006), (Kauppinen 2007), (Levy 2004), (Ludwig 2007)

> It is not especially clear why Mallon *et al.* think that the intuitions of undergraduate students might reveal something important about reference. The typical undergraduate student has no special training in semantics and so should not be expected to possess any intuitive insight about the nature of reference. For insight on the theory of reference, it makes much better sense to turn to smart, well-trained philosophers of language, such as Kripke and Evans. The undergraduates Mallon *et al.* surveyed are competent speakers of English (English is the language of instruction at HKU), that's true; but why suppose that mere competence suffices for intuitive insight? If one is competent in English, one can express one's beliefs and desires in English and be understood by other speakers of English. But being able to achieve these communicative goals is a far cry from knowing, even implicitly, how the reference of one's terms is secured. Knowing how the reference of one's terms is secured takes hard thinking and detailed semantic analysis and theorizing. There is no reason to think that every competent speaker is suited to this task, and certainly no reason to think that mere competence makes them suited to it.

Special training and the skill to use philosophical methods (from "hard thinking" to "detailed semantic analysis") are necessary to reach philosophically relevant results. This is what differentiates folk from experts. The methods of philosophers also include philosophical discussions and Levy points out that philosophy is a "distributed enterprise", where the intuitions and arguments (from intuitions) can be and have been discussed, criticized and improved by many philosophers over time. As even the disagreement by one individual philosopher is irrelevant to the "prevailing body of arguments", the disagreement of the folk, who lack the necessary skills to offer to the discussion the same level of relevance and quality as philosophers, is rather unimportant (Levy 2004):

> Since it is the intuitions of the community of experts, intuitions driven by reflection upon a wide range of analogous cases and upon confrontation with the best arguments available, that matter, the failure of the folk to share our views is simply irrelevant. We should no more worry what they think about our analyses than physicists should care about folk conceptions of folk and time.

Not all philosophers agree that while holding philosophers' intuitions in high regard, the folk intuitions should be completely dismissed as these (can) still have some relevance to the philosophical theories.

We can adopt a more charitable position and consider folk intuitions as a relevant part in the discussions about the theory of reference. It is essential, we might even say, for philosophy

to take folk intuitions into account: "Indeed, why should anybody care about what philosophers do if they just argued about their own inventions?" (Kauppinen 2007, p. 96). This does not necessarily grant folk intuitions an equal status with philosophers' intuitions. The expert intuitions have several advantages – they (a) have been drilled by relevant scientific theories, making them more at home at identifying and operating with the concepts in question, (b) are more likely to turn out to be correct when tested, and (c) help us to advance with our theories quicker, compared to folk intuitions. Because of these reasons, one would still prefer, if possible, the intuitions of philosophers, as the expertise in the field is important. It must be noted, that this expertise is not infallible. The fact that a person can operate successfully with a concept does not mean they are correct about that concept. Devitt argues that a person can have a competence in using a concept, but still remain *ignorant* about it (2006, p. 106):

> She is privileged in *her ready access to data*, not in *the conclusions she draws from* the data; conclusions of the competent, just like of the incompetent, are empirical responses to the phenomena and open to question; they arise from the empirical observation of the data.

The second reason for granting philosopher's intuitions a higher value is the methodological issues with the surveys conducted to research folk intuitions. Several philosophers have argued that the experimental philosophers' surveys fail to provide us data about the type of folk intuitions, relevant to the theories of reference. According to Kauppinen (2007), all these experiments give us is the information about the *surface intuitions* of non-specialists. This is not the same way philosophers use intuitions. "Appeals to intuition are not appeals to gut reactions,", as he puts it (p. 104).

Following two formulations show how "(I)" – the appeal to intuitions is usually constructed and "(E)" – how the experimentalists, according to Kauppinen, rephrase this (p. 98):

> (I) S; In S, we would (not) say that X is C

> (E) 'In S, we would (not) say that X is C' is a prediction that (most) non-specialists will (not) say that X is C if the case S is presented to them
> But, (E), when made more explicit, would actually be saying something which is quite irrelevant and does not answer the questions experimental philosophers want to ask (p. 105):

> (E*) 'In S, we would say that X is C' is a prediction that (most) non-specialists who (1') appear to understand the question will say that X is C if the case S is presented to them (2') however they consider it in whatever conditions they find themselves in and (3') whatever kind of considerations influence their response.

He sees these explications (both the simple (E) and more explicit (E*)) problematic, because of the embedded points of failure – of competence, performance and influence of irrelevant factors. He proposes an alternative (p. 101), conditions of which, if accepted, would prove the experimental philosopher's methods of using surveys incapable of harvesting the philosophically-significant *robust intuitions*:

> (A) 'In S, we would say that X is C' is a hypothesis about how (1) competent users of the concepts in question would respond if (2) they considered the case in sufficiently ideal conditions and (3) their answer was influenced only by semantic considerations.

Kauppinen claims that normal speakers of a language can manage using the concepts in everyday situations, but this does not equal to having the understanding of and using the concepts in a normatively correct way. He argues that for something to *seem right* and *be right* normative constraints are required; otherwise the concepts might lose their content altogether – when you cannot restrict the application of a concept according to some rules, it can be applied to everything (or nothing) and thus become unusable. Therefore, if we would just take all possible applications of a concept into account, we could not specify when a mistake is made, leaving these applications all equally correct, or incorrect. To differentiate the correct and incorrect uses of a concept, even by competent users who are less likely to make mistakes but still might, or by the majority of speakers who are actually mistaken in applying the concept, we need to specify the conditions under which the correct use of a concept would normally occur, and make sure the concept is treated semantically, not pragmatically.

The surveys, as he argues, might contain self-refuting elements. When control question are used in surveys to rule out mistakes or lessen any external effects which could influence the responses, it already presupposes that some uses of a particular concept are less correct. How do we know, which out of all possible uses is the correct one and actually does reflect the folk concepts accurately? To use surveys to find this out would be infinitively regressive

and to use intuition-based judgements would refute the Sceptical Experimental Philosophy. The experiments have to give us information about robust intuitions, and for this, they have to have the component of dialogue – an activity where the philosopher and the test subject are discussing and clarifying the responses. Therefore, according to Kauppinen, surveys are not a suitable way to find out about folk concepts. This is not an issue for finding a correct theory of reference, though. Having normative knowledge about the concepts from past experiences in the language game, philosophers are suitable to reflect on the concepts from the armchair – they have the necessary data and the tools for interpretation.

Kauppinen goes into great lengths to defend the armchair-type of philosophy, but does not succeed in all accounts. There are several assumptions in his arguments, which he takes for granted. Let's look at the following citation, where he explains how traditional (analytic) philosophers build their argumentation (p. 101):

> It should be obvious that when philosophers appeal to 'us' in making their claims, the extension is limited to those who are competent with the concept in question. After all, what incompetent users of a concept say about a given case does not tell us anything about the concept we are interested in—someone who has no relevant pre-theoretical knowledge about the concept cannot manifest it.

The general idea of only competent users of concepts to participate in the debate is reasonable. To read this in the context of his article, it is clear the competent people, when we are talking about philosophical concepts, are philosophers, not laypeople (i.e. non-philosophers). He does not think non-philosophers have the necessary pre-theoretical knowledge about the concepts in question. It is not clear though, why he thinks the concepts being surveyed by experimental philosophers are "technical", in the sense that their content cannot be grasped by non-philosophers. He says we would not ask a small child to reflect on a Gettier case – true, we would not. The main reason perhaps is that small children have not yet developed the counterfactual reasoning abilities, i.e. they are *unable* to reflect on Gettier cases. Competent adult speakers of a language, on the other hand, are sufficiently equipped to understand and discuss Gettier cases, even while lacking any special philosophical training.

What if these adult speakers use concepts differently? Kauppinen claims, that as the users of the (same) language, we ought to share the robust intuitions about concepts to make sense of each other (p. 110):

> Since we use language to communicate with each other and sharing concepts is necessary for agreement and disagreement, there is strong a priori reason to believe that people's robust intuitions will line up with each other, at least in central cases.

There has never been as much communication between different cultures from all around the world as there is now. The cultural variations in using concepts have perhaps only recently started to play as significant philosophical role for humankind. The robust intuitions, contrary to what Kauppinen argues, do not necessarily have to line up – because the communication between cultures does not necessarily work the same way anymore. Historical, cultural, conceptual disagreements between speakers of the same or of various languages do exist. The studies by Nisbett *et al.* have shown these kinds of variations, and this has fuelled the research by Machery *et al.*, and other experimental philosophers focusing more on the Western and East Asian comparative surveys. The possibility for conceptual variations or disagreements is very clearly one of the findings of experimental philosophy.

Kauppinen does argue that there can be an argument from disagreement, where people have different content for concepts (p. 108), but in this case they would be speaking about different things, or speaking "past each other", without any genuine disagreement about the concept in question. If the concept of "fast" contains the description X for one person and Y for another, then these are essentially different concepts, even though sharing the same name.

This position, I argue, would actually nullify any possibilities for finding different world views from our own – if two groups have a fundamentally different content for a concept Z, and there can be no disagreement as the concepts in this case, according to Kauppinen, would actually be different, then indeed, there can be no non-line-up between the concepts. At the same time, we would be unable to learn, as anything different from what we know would be labelled as a different concept. If what Westerners consider "duty" is what the law enforces and for East Asians it is something which is honourable, then are we dealing with

two distinct concepts, accidentally sharing a name, or is it still the same concept, with different content for different cultural groups?

Schroeter (2008), a "foe of conceptual analysis" as he describes himself, thinks that although folk intuitions can have a limiting influence on philosophical conceptual analysis because of its fundamental nature (i.e. being conceptual), they do not tie down other types of philosophy. The conceptual theories have to take robust intuitions of the folk into account and make sure the theory is in line with these. Non-conceptual theories, on the other hand, still have making sense of linguistic practice as a priority and the robust intuitions can be used as something to contemplate about, but not something to limit the results of the theory. Folk concepts can easily be dismissed, if they conflict with the actual linguistic practice. Experimental philosophers are identifying themselves more with conceptual analysts, but this might lead them to miss other relevant aspects of folk concepts – the non-criterial social content of the concepts. What a concept means, in a dictionary-sense, and what associations are made with the concept to meet the requirements for to be "that concept" can be different (he quotes, as an example, feminist research by Haslanger[16], regarding the concept of "woman" and its non-criterial associations, such as specific social roles, in the practice of using the concept).

Ludwig (2007), while describing how thought experiments, which's goal is to "to draw out the implicit knowledge we have of the application conditions of our concepts as it is embodied in our dispositions to deploy words expressing them", should be conducted, stresses that both the experimenter and the subject of the experiment should know how thought experiments work and what is the goal of that specific thought experiment.[17] If these demands are not met and the possible (and very likely) distractive factors – such as background beliefs, scenario comprehension, motivation, or plain mistakes – are not taken into account, the thought experiment, in a way, fails.

---

[16] Haslanger, Sally. 2000. Gender and race: (What) are they? (What) do we want them to be? Nous 43: 31–55.
[17] Ludwig has identified several possible problems if this is not so: when the subject does not know what is required from her, the response could be about what she thinks people would say; or she might fail to imagine herself in the described situation thus nor connecting with the story; or fail to identify the possible biases affecting the judgement; or the answer can be a from-the-top-of-her-head type, not really based on what content the concepts have, if consulted.

He uses an example from mathematics (infinite number series) and concludes that an experiment asking for folk answers might easily lead to results which show the folk have difficulties in applying the mathematical concepts (i.e. giving theory-wise wrong answers), but this does not mean the mathematicians are using the concept of "number" from the folk. To answer to a question in this type of an experiment requires the subject to have previous training, in the sense of lots of practice and reflections, in the domain of mathematics. The intuitions from untrained folk are of no interest. The same applies to complex topics in philosophy, such as the semantics of proper names, etc. It is not that folk concepts are irrelevant *per se*, but those of untrained folk certainly are. Philosophers, on the other hand, have all the necessary background and training and are in a better position to run the first-person thought experiments and arrive at relevant intuitive results. The value of the surveys is somewhat similar to what Schroeter sees as the role of folk intuitions to non-conceptual philosophy. They are something to consult – in this case, as ideas from someone who has not been involved in the practice of philosophy, and therefore potentially interesting, though absolutely non-committing. That is, it seems, surveys serve a rather anthropological goal for Ludwig, but to be usable as such, they need to be improved to remove the methodological issues described above.

But, if Schroeter is correct and analytic philosophy includes robust intuitions of the folk as a necessary component, we arrive at a question about the similarity of the controlled-domain mathematics and folk-related-domain of semantics. As Ludwig is building his argumentation on the analogy of folk experiments in the domain of mathematics, I believe he is mistaken in comparing this to philosophical thought experiments about, let's say, intentionality or the theory of reference. To ask the folk a mathematical question which requires knowledge of the well-fixed and documented structure of mathematics, is one thing. Experiments can yield correct and incorrect answers and the folk misunderstanding of a concept does not tell us anything about the content or correctness of the concept, but rather about the need to perhaps improve the Public Understanding of Science program. Asking the folk a question of a philosophical domain – such as about the concept of "intentional" – is something very different and does not have the luxury of comparing the answers to a well-established (analytic) framework, deeming the answer either correct or

incorrect. The content of the concepts in the philosophical or psychological domain is not fixed and does not exist without the folk, compared to mathematics which would have comprehensible value as a mind exercise (to an off-world explorer) even if all humans were dead. The answers to complex mathematical questions can be intuition-based, of course, but they can be checked, eventually, against the pre-set analytical framework (of rules) of mathematics. It is not clear if and how something similar can be done with philosophical concepts – what constitutes as a framework in this case? Devitt (forthcoming) of course thinks these two – mathematics and philosophy – do not differ much in the sense of justification and should have their concepts explained empirically.

One of the goals for experimental philosophers is to test if the content of the concepts, as understood and used by philosophers, actually holds. The clinically academic content of a concept could be false. If we take "false" as too strong a word, then even "does not hold or a significant part of competent speakers" would be sufficient to raise doubts about the correctness of the content according to philosophers. To motivate the relevance of folk understating of concepts, we need to make a distinction between technical language and ordinary language.

When we discuss terms such as "*a priori*" or "epistemology" or even "conjunction", the potential difference between the content of these terms for philosophers and folk is of little significance. These terms are coined and (mainly) used by philosophers. They are of no relevance, we might say, to folk – the same way as folk understanding of these terms is of no relevance to philosophers. These terms would actually fall into the "own inventions" type of terms, described as irrelevant to non-philosophers by Kauppinen. For concepts with folk relevance – "reference", "duty", "intentionality", "free will", etc. – the folk understanding is of relevance. I argue that if a concept has a real life application, and by this I mean that competent (but not philosophically or technically trained) speakers of a language can answer the challenge "How would you use this word in a sentence", a part of the meaning of the concept, unless we subscribe to extreme externalism, becomes shared. Folk use can even change the initial meaning, as it was the case with Madagascar. For technical concepts, the change in the concept's content can also occur, usually through scientific discoveries or conceptual analysis. If we would remove the folk part of the meaning from a concept, we

would end up with a technical concept which would have a content which is *completely different* from the one what the folk are using[18]. This would lead to the aforementioned "argument from disagreement" according to Kauppinen. To make sure we are talking about the same concept, and hence resolve this situation, we can declare which "version" of the concept is used – either technical or folk. This is not usually done explicitly, though[19]. It seems to me that philosophers think of the concepts they use as technical, but these are not "purified" before use, therefore they might still be using the concepts with elements of folk meaning. Philosophers, as folk, also have background beliefs. An infallible philosopher would be able to remove all background beliefs and any other unwanted variation[20] before sitting in an armchair to contemplate about a concept. It is difficult to prove, if philosophers operating e.g. with the concept "reference" have managed to do so – and the studies in experimental philosophy argue this has not been done successfully.

This all is not to say folk understanding *determines* the content for these concepts. The concept of "intentional" could really have more to it (i.e. additional folk meaning) than philosophers have realized, as the harm/help cases in Knobe's research suggest. We cannot yet be sure that the analysis on the data has yielded meaningful and philosophically significant results, as the methodological criticism towards experimental philosopher's studies holds in lots of aspects, but it is too early to dismiss these kinds of studies, at least based on the assumption that "laypeople know nothing".

I believe Kauppinen sees the experimental philosophy somewhat too narrowly – being only quasi-observational, conducting hands-off surveys and rather error-prone. As I have suggested, interdisciplinary approach to (experimental) philosophical research can help to iron out some, if not most, of the shortcomings of the current methodology.

---

[18] Wittgenstein would probably agree (Wittgenstein, L. (1958). Philosophical Investigations)

[19] Kuhn, when discussing incommensurability, doesn't seem to believe it even could be done (Kuhn, T. S. (1962). The Structure of Scientific Revolutions).

[20] See Chapter 5 for a discussion about unwanted variation

5. Methodological pitfalls and improvement options

Experimental philosophers claim to have used various scientific methods in their research about intuitions, from experimental psychology, cross-cultural psychology, social sciences, cognitive sciences, etc. This is in contrast with the traditional analytic epistemology projects, generally titled as "Intuition Driven Romanticism" by Weinberg *et al.* (2001):

(i)   The strategy must take epistemic intuitions as data or input. (It can also exploit various other sorts of data.)

(ii)   It must produce, as output, explicitly or implicitly normative claims or principles about matters epistemic. Explicitly normative claims include regulative claims about how we ought to go about the business of belief formation, claims about the relative merits of various strategies for belief formation, and evaluative claims about the merits of various epistemic situations. Implicitly normative claims include claims to the effect that one or another process of belief formation leads to justified beliefs or to real knowledge or that a doxastic structure of a certain kind amounts to real knowledge.

(iii)   The output of the strategy must depend, in part, on the epistemic intuitions it takes as input. If provided with significantly different intuitions, the strategy must yield significantly different output.

Cullen (forthcoming) acknowledges this, but claims the hoped-to-be-better (scientific) methodology as used by the experimental philosophers has its own issues. He describes the core of the methodology as *Survey-Driven Romanticism*: "People's philosophical intuitions are implanted within them in some way, and by administering simple surveys we can discover them". He does not think this is a successful enterprise, as the surveys conducted by experimental philosophers cannot be used to achieve their objectives of researching intuitions due to methodological shortcomings. In this chapter I will study more closely what kind of methodological criticism can be (and has been) put forward towards the

practices of experimental philosophers. Most of these pitfalls can be avoided by improving the experimental research methodology, and where applicable, I provide my recommendations for Significantly Improved Methodology for Experimental Philosophy (SIMEP). Some of the known methodological issues have been taken into account in subsequent studies, and the new experiments have been redesigned to avoid the aforementioned issues. This has not been done in regard to all issues, though, and from a rather "armchairy" position I raise some additional potential methodological issues, the existence and relevance of which should be empirically tested.

Before drawing any philosophically significant conclusions from experimental data, we have to question of this data is indeed about what we set out to study. Experiments are often conducted in artificial settings and use thought experiments instead of observing the real-life situations. Questions, asked under these conditions, can be interpreted in a non-uniform way and the answers are not necessarily about the same question as the experimenters had in mind. If one is asked, in a controlled environment, about their intuitions on reference with thought experiments using unfamiliar names and situations, the response is not necessarily about what the experiment's participant's intuitions say about reference in everyday situations, but about their theoretical view on the correct theory of reference. The participant has time and can (or might even have to, in case the probe's story seems strange) analyze the situation in ways very different from everyday judgements. This situation is similar to what happens in anthropology, if the observation is replaced by direct involvement and influence. Therefore the answer does not give much insight into intuitions about reference. The variation in the supported theories of reference, which is the actual gathered data, is common even among philosophers, as Marti (2009) notes. More "anthropological" approach to the experiments can potentially solve this issue. Assuming that laypeople's intuitions are relevant for theories of reference, instead of classroom-setting research conducted among undergraduates, random "bus stop surveys", although seemingly anecdotal, are more likely to give less over-analyzed answers and portray the intuitions (about reference) more in line with the actual use of reference, instead of theoretical reflections. I hence suggest the first rule for SIMEP:

> (SIMEP-1)   Experiments, which are meant to elicit laypeople's intuitions about the *actual use* of reference or any other applicable concept, must be conducted in a maximally real-life-like environment using probes with maximally familiar situations.

Once we have decreased the likelihood of asking about the wrong type of intuitions or actually not asking about intuitions at all, we should make sure the participants understand the probes and questions used in the experiments. Machery *et al.*, but also several other experimental philosophers, have ran their surveys with both Western (American) and Chinese participants using probes in English. It has been possible because the popular choice for a venue from East Asia has been the University of Hong Kong, which is an English-language based university, and its undergraduates are claimed to be fluent in English. Research in philosophy of language has shown that some philosophical problems disappear, when the discussion is transferred to another language. It is highly unlikely the average proficiency in English as a second language is equal to the average proficiency of English as a first language among undergraduates of respective universities. This means the Chinese participants are using an additional "interpretation" layer when reasoning about the probes. This layer of linguistic competence can affect both the understanding of the probes and the understanding of the questions – instead of immediate intuitions the responses can be about the suspected application of reference in a foreign (i.e. English) language. Barry Lam (2010) describes an experiment with Chinese participants, using probes similar to those of Machery *et al.*, but in Cantonese. The findings of this survey do not support Machery *et al.*'s findings – according to Lam, Chinese are not actually the supporters of the descriptive theory of reference, as the survey shows their responses to be in line with the causal-historical theory of reference. It is unclear if the body of research which considers the English-speaking undergraduates from Hong Kong to be a sample of East Asians suitable for intuition diversity probing has settled for this out of convenience, as in this case Cantonese language skills are not required from the researchers, or because of a genuine belief about the suitability of this group of participants. As Lam's experiment shows, this position can be mistaken, and native language probes will result in different results from probes in one's second language. If the probes can be misinterpreted (from the

point of view of the experimenters), then it is not safe to assume the instructions for the experiment are free from a similar effect. These issues lead us to the second rule:

(SIMEP-2)   Experiments must be conducted in the participant's native language; this applies to both the experiment's instructions and the probes.

Some misinterpretation can occur also in one's native language due to the ambiguity of the probes, an issue that has received a lot of attention from the (methodological) critics of experimental philosophy[21]. It has been suggested that using e.g. (a) quotes, (b) intensifying adverbs, (c) contrasting phrases and (d) colloquial phrases, can create confusion about the meaning of the probe texts and the attached questions. The level of likelihood for these factors to actually influence the response is difficult to measure, as it depends on both the linguistic competence and the personal cognitive processes (i.e. how much one can "read into" specific phrases used). Its avoidance is therefore also difficult to achieve, but best effort can be applied:

(SIMEP-3)   Experiments' phrasing must be cleared from as many ambiguous and confusing constructions as possible; thorough reflection by speakers with different socio-economic backgrounds to identify ambiguities is recommended.

A closely connected issue with the general ambiguity is the distinction between the speaker's and semantic reference. When presented with a question about the probe, should the participant answer it from the protagonist's perspective or from the external omniscient observer's, Pethö asks (Pethö 2005). Deutsch has argued (Deutsch 2009) that the pragmatic speaker's meaning, which seems for him to be more likely elicited by the experiment's phrasing, is actually irrelevant to the philosophical discussion (about the theory of reference). As with the above described ambiguity, it is difficult to say with an absolute certainty which interpretation is preferred by the participant. The gathered data includes, potentially, both interpretations and as the difference is philosophically significant, the data cannot be interpreted to show genuine disagreement between Western and East Asian participants regarding reference. Sytsma and Livengood run several experiments in classroom and non-classroom settings and have reported their findings (Sytsma &

---

[21] See e.g. (Cullen forthcoming), (Deutsch 2009), (Marti 2009), (Pethö 2005), (Sytsma & Livengood forthcoming)

Livengood forthcoming) that different interpretations of the perspective do indeed influence the answers – something that should be avoided. Some variation still remains in the answers, even when the expected perspective is stated more clearly. The authors believe that a 80% - 90% agreement is sufficient to count as consensus and a 100% agreement should not be pursued. The reasons for this are possible measurement noise and some actual variation of intuitions, both of which would carry insignificant weight for the overall results.

(SIMEP-4)   Questions about the probes have to make it explicit which perspective the participant is expected to take.

Even when the perspective is clear, the participant can have an intuition to answer something else than one of the two options usually provided. A third option might be required – not an option on the same level as the "default" options, but a more generic one – such as "none of the above, but...". This leaves room for different interpretations – including those the study conductors are not able to foresee. These kinds of open questions, of course, make the survey analysis a lot more complex. Dichotomous closed questions in general can force participants to adopt extreme positions, which would not be their choice if the questions would be of the open type. An alternative third option is just to state that the probe does not have enough information to make an informed choice between the first two options. [22]

(SIMEP-5)   Open non-dichotomous questions should be preferred when designing the probes for experiments. In case dichotomous questions have to be used, an "other" option has to be included.

As I've shown above, it has been argued that competency with the terms in question is necessary for harvesting any relevant data for intuitions about those terms. In addition for the requirement for conceptual competency – questions e.g. about string theory do not elicit any meaningful intuitions from the folk – the context competency is also required. A point I would like to make – one that definitely requires structural empirical research for validation – is that some specific contexts of the experimental philosopher's probes are unsuitable for harvesting meaningful data. Having discussed several of the latest experiments, dealing

---

[22] See e.g. (Pethö 2005), (Sosa 2009) and (Cullen forthcoming)

with morality and intentional action not reference, with people in my closer circles, I've become sceptical of the suitability of military- and management-related thought experiments for eliciting meaningful folk intuitions. The concept of (morally) good and bad in these cases does not necessarily follow the folk understanding of these concepts. Management usually requires more information about potential side effects and viable alternatives before making a decision of implementing a new (environmentally-harmful) programme than a short message from the assistant. The decisions are not made hastily, but when a quick decision is needed, management often relies on "gut feeling" – a type of intuition which has taken many years to mature and contains more variables than a person from the street could ever imagine. If a layperson is asked about the intentionality of harming the environment, the best answer one can give is "in the CEO's position, I would ..." but this is rather close to asking layperson a question from theoretical physics or economics and hoping for a meaningful insight. I subscribe to the argument that one has to be competent in the concepts one is asked about, and string theory is not that much different a topic from any in the management realm in this aspect. Similar effects can be noted in military-related cases, where the decisions – either long-weighed or immediate – are often tactical, not pragmatic in layperson's terms. For a platoon leader to send his men into certain death can seem like an extremely stupid, short-sighted and cynical decision, but the larger tactical picture of the situation is unknown and perhaps even unimaginable for a layperson. These two types of thought experiments have produced interesting results, but the framing effects in this case could be too strong to use any of the gathered data for any kind of analysis.

(SIMEP-6)   Probes should not use concepts or contexts for which the participants lack the competence.

To decrease the effects of observer or subject bias, improved methods from other areas of science can be used. One option would be to use *blind methods*, either single-blind or for added efficiency, a double-blind method, also suggested by Weinberg (2007) and (2009). According to this method, the people actually conducting the research do not know, before the data is gathered, what and how exactly is studied and which subjects belong to the experimental group and which ones to the control group.

Karu, Kaimar
Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?

(SIMEP-7)   To avoid the experiments' contamination by bias, single-blind methods minimum, and double-blind methods preferably must be used.

6.   Subsequent studies in experimental philosophy

Since Machery *et al.*'s experiments described in (Machery et al. 2004), and the following criticism on the methods employed by experimental philosophers, various more advanced in experimental philosophy have been conducted. From these studies, two main lines have emerged. The first are the experiments on the "Knobe effect", where folk understanding of intentional action is studied[23]. The second batch is the follow-up experiments on cross-cultural variation, either to strengthen Machery *et al.*'s initial findings or refute them. In combination with the relevant criticism, Machery *et al.*'s conclusions on their gathered data are indeed under serious scrutiny.

Barry Lam (2010) believes that while Machery *et al.* criticized Kripke for philosophizing out of an armchair and thus not considering the intuitions of other (non-Western) people, they have made a similar mistake by presenting the experiments' vignettes to East Asian participants in English and with English names for the agents, not in Cantonese, leaving room for linguistic competence related differences, as opposed to intuition-related differences reported by Machery *et al.* He conducted an experiment (with 38 Cantonese-speaking and 31 English-speaking participants of various socio-economic statuses) modelled on the Gödel case, replacing the name and context of "Gödel" with that of "Shakespeare", and presented to participants in English and Cantonese respectively. The results show that both Cantonese and English-speaking participants had answers in line with causal-historical theory of reference. In another experiment (with 33 Cantonese-speaking and 34 English-speaking participants, similar to those in the 1st experiment), this time Julius-type stories instead of Gödel-type, he used two stories. First story used made-up names, this time

---

[23] E.g. (Knobe 2006), (Machery 2008), (Mallon 2008), (Phelan & Sarkissian 2009), (Pettit & Knobe 2009)

different in English and Cantonese, and also with a different context, more in line with the participant's background (Western vs. Chinese story). Second story was a modification of the story from the first experiment, rendered to be Julius-type. The results showed no significant difference between the reference use for English-speaking and Cantonese-speaking participants. The results of these experiments show that whatever the reasons for the differences shown in Machery *et al.*'s study, they are influenced by the language, not by the difference in prevailing theory of reference or intuitions about it.

Marti (2009) has argued that the original experiments by Machery *et al.* (2004) asked wrong questions from the participants and therefore elicited only metalinguistic intuitions (about theories of reference) from the participants, as opposed to linguistic intuitions (about the use of names) which should be of relevance to the theory of reference. Her recommendation for future experiments was (a) to distinguish between observations that will reveal how people *actually* do things and how they *think* they do them or what they *think* is correct, (b) to "flesh out" the test, i.e. make it clear if the participants have background knowledge about the persons in the stories and if the story is supposed to be counterfactual or of pretend actual world and (c) use open questions instead of dichotomous options.

Machery *et al.*'s (2009) response was that Marti's distinction between metalinguistic and linguistic intuitions is indeed correct, but there is no reason to believe these intuitions should differ in content. If that would be true, the authors claim, the philosophy of language is in trouble because it has ever since used cases to elicit intuitions which could in fact be philosophically irrelevant metalinguistic intuitions.

To prove that both types of intuitions are similar and their initial study is still relevant, they conducted a new variation of their original experiment. It includes two probes (or vignettes), one with a linguistic and one with a metalinguistic question. They did not take Marti's advice to include open questions in their experiment, though, explaining that it would not produce easily analyzable data.

> Probe: Ivy is a high school student in Hong Kong. In her astronomy class, she was taught that Tsu Ch'ung Chih was the man who first determined the precise time of the summer and winter solstices. But, like all her classmates, this is the only thing she has heard about Tsu Ch'ung Chih.

> Now suppose that Tsu Ch'ung Chih did not really make this discovery. He stole it from an astronomer who died soon after making the discovery. But the theft remained entirely undetected and Tsu Ch'ung Chih became famous for the discovery of the precise times of the solstices. Everybody is like Ivy in this respect; the claim that Tsu Ch'ung Chih determined the solstice times is the only thing people have heard about him.

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

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

The experiment (run in 3 countries – India, France and Mongolia – with a total of 227 participants) shows similar responses to both linguistic and metalinguistic probes with no significant variation between the three countries, but with significant intra-cultural variation. The results show that linguistic and metalinguistic intuitions elicit largely the same answers, thus rendering Marti's criticism noteworthy, but insignificant to their arguments.

The authors readily provide three responses to possible criticism than can be targeted at their new experiment's successfulness to elicit linguistic intuitions, as the name of the referent in their story is mentioned, not used. They claim that although the name was mentioned, not used, it is in line with Marti's proposed improved experiment. Secondly, the assessment of a sentence's truth value (as in "Do you think her claim is true") as a response shows what the participant would say. Thirdly, this method is common to semantics.

Characteristically to experiments with Machery's involvement, the applied methodology is loaned from another science (in this case, semantics) or from another branch of philosophy (philosophy of language) and defended on the basis of its usage and trustworthiness there. Any criticism towards their methodology would in this case target the source science as well as their experiment. Borrowing a methodology does not guarantee its correct application, though. It is unclear from the article if and how the truth value assessment of a sentence in

case of reference enquiry works similar to how the participant would actually use the reference. By designing new questions for their probe, they claim to have followed Marti's suggestions, but have done so only partially. An important distinction between their phrasing of the question and that of Marti's is the open/dichotomous answer option. It is quite clear that open questions are more difficult to analyze, but to achieve what Marti proposed – to understand how reference is actually used – the effort must be made. Pethö (2005), Sosa (2009) and Cullen (forthcoming) all support the importance of revising the answer options to harvest more precise data and identify possible misunderstandings and unwanted variation in the experiments. Machery *et al.* have not shown convincingly that their modified probe is in fact able to elicit linguistic intuitions, as defined by Marti, instead of the potentially irrelevant metalinguistic intuitions. If the probe and the attached questions can be interpreted in more than one way, the value of their alleged findings of intra-cultural variation remains a weak hypothesis.

The recent study by Sytsma and Livengood (forthcoming), already briefly described in the methodological criticism section, set out to clarify the epistemic perspective ambiguity in the original study by Machery *et al.*, i.e. to design experiments where it is clear whose perspective is expected to be applied by the participant, and to show that the original study was ambiguous.

In their first experiment with 189 Western participants from the University of Pittsburgh, the questions were rephrased to clarify whose perspective is expected to be applied (p. 11):

> **Original**: When John uses the name "Gödel," is he talking about: (A) the person who really discovered the incompleteness of arithmetic? Or, (B) the person who got hold of the manuscript and claimed credit for the work?

> **John's Perspective**: When John uses the name "Gödel," does John think he is talking about: (A) the person who the story says really discovered the incompleteness of arithmetic? Or, (B) the person who the story says got hold of the manuscript and claimed credit for the work?

> **Narrator's Perspective**: When John uses the name "Gödel," is he actually talking about: (A) the person who the story says really discovered the incompleteness of arithmetic? Or, (B) the person who the story says got hold of the manuscript and claimed credit for the work?

The number of (B) answers for different perspectives was as follows: original 39.4%, John's perspective 22.0%, and narrator's perspective 57.4%. John's perspective seems to elicit descriptive intuitions, whereas narrator's perspective elicits causal-historical intuitions. As the results for the narrator's probe still varied significantly, the probe with an updated version of the question was phrased and presented in a second experiment (p. 14):

> **Clarified Narrator's Perspective**: Having read the above story and accepting that it is true, when John uses the name "Gödel," would you take him to actually be talking about: (A) the person who (unbeknownst to John) really discovered the incompleteness of arithmetic? Or, (B) the person who is widely believed to have discovered the incompleteness of arithmetic, but actually got hold of the manuscript and claimed credit for the work?

The results of the second experiment show that (the combined average) 74.9% of participants (from total 142) chose answer (B), which is in line with the causal-historical theory of reference. It is noteworthy to mention that although this experiment included participants, among others, who had been classified as philosophers[24], there was no statistically significant difference between their answers and those of non-philosophers.

The third experiment, which was run face-to-face in a non-classroom setting with all of the four probes (with additional filler probes) presented to all participants, repeated the combined results for the first and second study quite closely – 74.3% of participants (from total 35) chose answer (B) for the clarified narrator's perspective question.

In the fourth experiment, the participants were presented with the original Gödel probe from Machery *et al.*, and were then asked which of the following restatements best corresponds with their understanding of the test question:

> (1) When John uses the name "Gödel," does John think he is talking about: (A) or (B). Or, (2) When John uses the name "Gödel," is he actually talking about: (A) or (B).

---

[24] "Participants were classified as philosophers if they were a professor of philosophy, had completed (or were in the process of completing) a graduate degree in philosophy, or had completed (or were in the process of completing) an undergraduate major in philosophy." (p. 15 fn 10)

The results show that from participants answering (A) the majority understands the question from John's perspective, while from participants answering (B) a slight majority understands the question from narrator's perspective.

Based on their results, Sytsma and Livengood conclude that Machery *et al.*'s instrument is broken and therefore does not provide any significant (usable) data for the analysis of people's semantic intuitions, which is why their experiment is not successful in showing cross-cultural or intra-cultural variation of intuitions about reference. Some variation still remains even in the improved experiments answers. The authors believe that a 80% - 90% agreement is sufficient to count as consensus and a 100% agreement should not be pursued. The reasons for this are possible measurement noise and some actual (individual) variation of intuitions, both of which would carry insignificant weight for the overall results.

As a) the instrument is deemed broken, b) the newly gathered experimental results do not support Machery *et al.*'s results for Westerners and c) the failure of the probe to reliably test semantic intuitions in case of Westerners casts a dark shadow over the instrument's ability to succeed in case of East Asians. The authors conclude that as the same methodology was used to test East Asian subjects, the findings – whatever they are – give us corrupt and un-interpretable data.

7. Discussion

In addition to improving experimental philosophers' research by taking into account the methodological criticism described in Chapter 5, there are some fundamental changes – hopefully improvements – that could be done to the experimental philosopher's approach.

At first, we should accept that the intuitions, either used by philosophers in their armchair-philosophizing and philosophical debates, or those of folk, do not necessarily provide us with truth. As Williamson and Devitt both have noted, intuitions are a good way to gather data for the advancement of current theories, but not a waterproof reliable source of knowledge. Philosophers have used intuitions in their practice for a long time, and there is nothing wrong with it, or anything shameful about it. The mistake is made when intuitions are taken to provide us with *a priori* data, or when philosophers' intuitions are taken for granted as being superior to folk intuitions. Devitt argues that expert's intuitions are more likely to turn out to be true when tested and I agree with his argument. I believe one has to be cautious, though, to acknowledge the limits of available testing methodologies to verify or refute any non-conformist intuitions. Intuitions about a concept which are not in line with what we consider the concept to be can be missed or deemed irrelevant.

Secondly, the criticism towards experimental philosophy's methodology needs to be taken seriously for any new experiments to harvest relevant data and for the analysis to produce relevant (and solid) information, instead of hypotheses on shaky grounds. Distinction between the semantic and pragmatic use of the terms in the probes has to be addressed, so we know if the elicited intuitions are about theoretical approach to the questions, or about the actual use of the concepts discussed. As for the probes' language, every detail of the phrasing must be scrutinized to exclude all potential misunderstandings from the text. The more room there is for interpretation of what the described scenarios say, what the

questions ask about or what the experiment's authors expect the participants to answer, the less reliable data we get.

For cross-cultural experiments the language issue is perhaps the most pressing one – using probes in English to ask East Asian participants about their intuitions, even if the participants are judged to be "fluent in English", does not necessarily elicit "pure" intuitions. The participants can either not understand the probes and/or the questions therein, or they answer the questions from a metalinguistic position – how they understand the concepts are used in English, not how they would use the concepts in their native language.

The improvements about the probes' language might make it easier for the participants to choose (only) one correct answer from the set provided in the probe, but this does not mean that a limited set of answers is the best approach. The earlier experiments were avoiding open questions because of the difficulties in following data analysis, as several authors have also admitted. Giving a limited set of answer options without a "none of the above, but ..." option, or "yes or no" options instead of a scaled answer option can lead to a very restricted set of responses, suitable perhaps for the support of an argument, but not for gathering accurate data about intuitions or for impartial analysis of the gathered data.

Participant's responses can also be forcefully (and unconsciously) framed by other factors than the probes' language or the answer options. Socio-cultural status is perhaps one of the most obvious effects on the participant's responses due to large differences in education and experience, but more subtle effects, such as the context of the probe's story, the order of the probes presented, or even the experiment's environmental settings can have a significant effect on the results. If these (potential) effects are not accounted for and eliminated from the experiments' settings, the importance of the resulting data can be intellectually interesting, but unsuitable for any further analysis.

The diapason of experimental research in philosophy has widened significantly during the years, but can still be difficult to connect to experiments conducted in natural sciences, clinical psychology, etc. Recently, there has been "more scientific" research on connection with people's judgements and their brain activity. As discussed in Chapter 3.6, the

intuitions, if we accept their relevance for philosophy, require a "solid" empirical foundation for naturalists. These interdisciplinary studies, combining philosophy with practical psychology, give us empirical proof about people's moral judgements and their brain activity. Participants' responses are still based on intuitions, but these responses can be measured even before they have entered the participant's consciousness, thereby significantly decreasing several framing effects and changing the discourse from "this is what we think how people reason" to "this is how people's reasoning works".

In addition to increasing the "natural" part of research in experimental philosophy, I believe a qualitative jump has to be made for experimental philosophy to remain relevant and escape the destructive course. So far, the experiments have been conducted mostly by Western philosophers with undergraduates in Western and East Asian universities. Two new directions should be pursued, I believe.

The first direction concerns experiments' participants. As with the probes' language, the majority of subjects are undergraduates probably not because the philosophers have considered their intuitions to be most relevant, but because undergraduates are one of the most accessible research subject groups in academic settings. The framing effects, as laid out in the methodological criticism section, are clearly noticeable in this selection. An undergraduate belongs to a different socio-economical group than, for instance, a farmer on a rice field or a railway worker, working on tracks and about to be hit by a wagon. Resulting from this, their intuitions about moral thought experiments might also differ. To gather a wider variety of intuitions, more of various socio-economical groups should be studied. In a way, the experimental philosophers should stop relying on semi-experimental research, based on most convenient options (undergraduates as subjects, closed questions in the probes). Although this research has already showed some philosophically interesting variation in intuitions, more is needed.

The second direction invites East Asian philosophers to join the debate. The participation of Barry Lam, is highly appreciated, as his background seems to give him a better access to Asian way of thinking. Input from more "real" Asian philosophers is needed to break the experimental philosophy free of a larger armchair of Western thinking, where it has fallen

after setting the analytic armchair in fire. Levi (2004) agrees with this position: "The problem is not that we have ignored too many of the non-Western folk, however, but that we have not brought nearly enough non-Western philosophers into our conversation." Before investing heavily in the improvements of experimental philosopher's methodology, we should inquire how the questions we are currently tackling under the experimental philosophy's flag are received by East Asians. This will give us more data on both the relevance of the questions and on how to better design the experiments to ask the right questions and elicit relevant intuitions from the East Asian subjects.

I hope the future studies in experimental philosophy will change their course and instead of attacking analytic philosophy by proposing that "it might not work" or defending analytic philosophy by attacking other experimental philosopher's experiments, the focus will be on actually making new discoveries. Aforementioned clinical psychology and both new directions I suggested would be a good start for this.

In the next step of my own future research I would like to design and run methodologically improved and participants-wise more heterogeneous studies with all the recommendations for SIMEP, as described in chapter 5, taken into account.

References

Alexander, J forthcoming, 'Is Experimental Philosophy Philosophically Significant?', *Philosophical Psychology*.

Alexander, J, Mallon, R & Weinberg, JM forthcoming, 'Accentuate the Negative', *Review of Philosophy and Psychology*.

Cullen, S forthcoming, 'Survey-Driven Romanticism', *Review of Philosophy and Psychology* , vol 1, no. 3-4.

Deutsch, M 2009, 'Experimental Philosophy and the Theory of Reference', *Mind & Language Vol. 24, No. 4*, pp. 445-466.

Deutsch, M 2010, 'Intuitions, Counter-Examples, and Experimental Philosophy', *Review of Philosophy and Psychology*, p. online.

Devitt, M 2006, *Ignorance of Language*, Oxford University Press.

Devitt, M 2008, *Experimental Semantics*, viewed 18 April 2010, <http://experimentalphilosophy.typepad.com/experimental_philosophy/files/Devitt-Experimental2.doc>.

Devitt, M forthcoming, 'No Place for the A Priori', in MJ Shaffer, M Weber (eds.), *New Views of the A Priori in Physical Theory*, Rodopi Press, viewed 18 April 2010, <http://web.gc.cuny.edu/Philosophy/faculty/devitt/NoPlace.doc>.

Fine, A 1975, 'How to Compare Theories: Reference and Change', *Nous*, vol 9, no. 1, pp. 17-32.

Frege, G 1892, 'On Sense and Reference (Über Sinn und Bedeutung)'.

Grice, HP 1991 (1989), *Studies in the way of words*, Harvard University Press.

Huebner, B, Bruno, M & Sarkissian, H 2009, 'What Does the Nation of China Think About Phenomenal States?', *Review of Philosophy and Psychology*, 2009.

Kauppinen, A 2007, 'The Rise and Fall of Experimental Philosophy', *Philosophical Explorations*, vol 10, no. 2, pp. 95-118.

Knobe, J 2006, 'The concept of intentional action: a case study in the uses of folk psychology', *Philosophical Studies*, vol 130, pp. 203-231.

Kripke, SA 1980, *Naming and Necessity*, Harvard University Press 2001, Cambridge, Massachusetts.

Lam, B 2010, 'Are Cantonese-speakers really descriptivists? Revisiting cross-cultural semantics', *Cognition*, no. forthcoming.

Levy, N 2004, *Experimental Philosophy: A Critique*, viewed 4 April 2010, <http://experimentalphilosophy.typepad.com/experimental_philosophy/files/experimental.pdf>.

Liao, MS 2008, 'A defense of intuitions', *Philosophical Studies*, vol 140, no. 2, pp. 247-262.

Ludwig, K 2007, 'The Epistemology of Thought Experiments: First Person versus Third Person Approaches', *Midwest Studies in Philosophy*, vol XXXI.

Machery, E 2008, 'The Folk Concept of Intentional Action: Philosophical and Experimental Issues', *Mind & Language*, vol 23, no. 2, pp. 165-189.

Machery, E, Mallon, R, Nichols, S & Stich, SP 2004, 'Semantics, cross-cultural style', *Cognition 92*, pp. B1-B12.

Machery, E, Olivola, CY & de Blanc, M 2009, 'Linguistic and metalinguistic intuitions in the philosophy of language', *Analysis*, vol 69, no. 4, pp. 689-694.

Mallon, R 2008, 'Knobe versus Machery: Testing the Trade-Off Hypothesis', *Mind & Language*, vol 23, no. 2, pp. 247 - 255.

Mallon, R, Machery, E, Nichols, S & Stich, S 2009, 'Against Arguments from Reference', *Philosophy and Phenomenological Research LXXIX No. 2*, pp. 332-356.

Marti, G 2009, 'Against semantic multi-culturalism', *Analysis*, pp. 42-48.

May, J, Sinnott-Armstrong, W, Hull, J & Zimmerman, A 2010, 'Practical Interests, Relevant Alternatives, and Knowledge Attributions: an Empirical Study', *Review of Philosophy and Psychology*.

Nadelhoffer, T & Feltz, A 2008, 'The Actor–Observer Bias and Moral Intuitions: Adding Fuel to Sinnott-Armstrong's Fire', *Neuroethics*, vol 1, no. 2, pp. 133-144.

Nadelhoffer, T & Nahmias, E 2007, 'The Past and Future of Experimental Philosophy', *Philosophical Explorations*, vol 10, no. 2, pp. 123-149.

Nisbett, R 2003, *The geography of thought: how Asians and Westerners think differently. and why*, Free Press, New York.

Pethö, G 2005, 'On intuitions about proper names', *Acta Linguistica Hungarica*, vol 52, no. 2-3, pp. 281-301.

Pettit, D & Knobe, J 2009, 'The Pervasive Impact of Moral Judgment', *Mind & Language*, vol 24, no. 5, pp. 586 - 604.

Phelan, M & Sarkissian, H 2009, 'Is the "Trade-off Hypothesis" Worth Trading For?', *Mind & Language*, vol 24, no. 2, pp. 164 - 180.

Schroeter, F 2008, 'Experimental philosophers, conceptual analysts, and the rest of us', *Philosophical Explorations*, vol 11, no. 2, p. 143–149.

Sinnott-Armstrong, W 2008, 'Framing Moral Intuitions', in W Sinnott-Armstrong (ed.), *Moral Psychology vol 2, The Cognitive Science of Morality*, MIT Press, Massachusetts.

Sosa, E 2008, 'Experimental Philosophy and Philosophical Intuition', in J Knobe, S Nichols (eds.), *Experimental Philosophy*, Oxford University Press, New York.

Karu, Kaimar
Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?

Sosa, E 2009, 'A Defense of the Use of Intuitions in Philosophy', in D Murphy, M Bishop (eds.), *Stich and his critics*, Blackwell Publishing.

Sytsma, J & Livengood, J forthcoming, 'A new perspective concerning experiments on semantic intuitions', *Australasian Journal of Philosophy*.

Weinberg, JM 2007, 'How to Challenge Intuitions Empirically Without Risking Scepticism', *Midwest Studies in Philosophy*, vol XXXI, pp. 318 - 343.

Weinberg, JM 2009, 'On doing better, experimental-style', *Philosophical Studies*, vol 145, no. 3, pp. 455-464.

Weinberg, JM, Nichols, S & Stich, S 2001, 'Normativity and Epistemic Intuitions', *Philosophical Topics*, vol 29, no. 1-2, pp. 429-460.

Williamson, T 2004, 'Philosophical 'Intuitions' and Scepticism about Judgement', *Dialectica*, vol 58, no. 1, pp. 109-153.

Wright, JC Forthcoming, 'On intuitional stability: The clear, the strong, and the paradigmatic', *Cognition*.

Karu, Kaimar
Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?

Summary

Title: **Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?**

Using intuitions as evidence in philosophy has been challenged, and as a response to the challenge it has been argued that in some disputable cases intuitions were not used at all. I show that intuitions have been used even in those cases in question, and that there is little reason to dismiss intuitions as evidence without further analysis. I argue that intuitions can provide us with relevant information and discuss the suitability and relevance of folk intuitions, compared to those of philosophers'. I argue that for some tasks philosophers' intuitions are of higher value, but for other tasks folk intuitions are of equal relevance, even if not of equal accuracy. Recent studies in experimental philosophy have shown some variance in intuitions of Westerners and East Asians – I describe the studies, analyze the methodological pitfalls and propose several improvements for future studies to gather relevant and accurate data about folk intuitions.

Karu, Kaimar

Are intuitions (of lay-speakers) relevant for determining which theory of reference is right?

Resümee

Pealkiri: **Kas (tavainimeste) intuitsioonid on õige osutusteooria leidmisel olulised?**

Intuitsioonide kasutamist filosoofilise argumentatsiooni tõendusmaterjalina on kritiseeritud ning ühe vastusena sellele kriitikale on väidetud, et konkreetsetes kõnealustes argumentatsioonides ei ole intuitsioone üldse kasutatud. Ma näitan, et intuitsioonid on nendel kordadel siiski kasutust leidnud, kuid et intuitsioonide ebakõlbulikuks tõendusmaterjaliks tunnistamine vajab edasist analüüsi. Ma väidan, et intuitsioonid suudavad pakkuda meile relevantset informatsiooni ning arutlen tavainimeste intuitsioonide sobivuse ja olulisuse üle võrreldes filosoofide intuitsioonidega. Ma väidan, et teatud ülesannete jaoks on filosoofide intuitsioonid kõrgema väärtusega, kuid teiste ülesannete jaoks on tavainimeste intuitsioonid samaväärsed, isegi kui mitte sama täpsed. Hiljutised uuringud eksperimentaalfilosoofias on näidanud Lääne ja Ida inimeste intuitsioonide erinevust – ma kirjeldan neid uuringuid, analüüsin metodoloogilisi nõrkusi ning teen mitmeid parendusettepanekuid edasiste uuringute tarbeks, mis muudaks võimalikuks tavainimeste intuitsioonide kohta relevantsete ja täpsete andmete kogumise.