THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia
ISSN: 0495-4548
theoria@ehu.es
Universidad del País Vasco/Euskal Herriko Unibertsitatea
España

Loaiza, Juan R .
Molyneux's Question in Berkeley's Theory of Vision
THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia, vol. 32, núm. 2, 2017, pp. 231-247
Universidad del País Vasco/Euskal Herriko Unibertsitatea
Donostia, España

Molyneux's Question in Berkeley's Theory of Vision*

Juan R. Loaiza

Received: 29/02/2016
Final Version: 06/04/2016

BIBLID 0495-4548(2017)32:2p.231-247
DOI: 10.1387/theoria.15984

ABSTRACT: I propose a reading of Berkeley's *Essay towards a New Theory of Vision* in which Molyneux-type questions are interpreted as thought experiments instead of arguments. First, I present the general argumentative strategy in the NTV, and provide grounds for the traditional reading. Second, I consider some roles of thought experiments, and classify Molyneux-type questions in the NTV as *constructive conjectural* thought experiments. Third, I argue that (i) there is no distinction between Weak and Strong Heterogeneity theses in the NTV; (ii) that Strong Heterogeneity is the basis of Berkeley's theory; and (iii) that Molyneux-type questions act as illustrations of Strong Heterogeneity.

Keywords: Berkeley, Molyneux's question, Heterogeneity, thought experiments, perception.

RESUMEN: Propongo una lectura del *Ensayo sobre una nueva teoría de la visión* de Berkeley en la que las preguntas tipo-Molyneux son interpretadas como experimentos mentales en lugar de argumentos. Primero, presento la estrategia argumentativa general en el NTV, y ofrezco motivos para la lectura tradicional. Segundo, atiendo a algunos roles de los experimentos mentales, y clasifico las preguntas tipo-Molyneux en el NTV como experimentos mentales *constructivos conjeturales*. Tercero, sostengo que (i) no hay distinción entre tesis de la Heterogeneidad Débil y Fuerte en el NTV; (ii) que la Heterogeneidad Fuerte es la base de la teoría de Berkeley; y (iii) que las preguntas tipo-Molyneux actúan como ilustraciones de la Heterogeneidad Fuerte.

Palabras clave: Berkeley, pregunta de Molyneux, Heterogeneidad, experimentos mentales, percepción.

1. Introduction

One of the main theses in Berkeley's *An Essay towards a New Theory of Vision* (1709)[1] is the claim that there is no necessary connection or common idea between our different sensory modalities. I will call this the Heterogeneity Thesis. In the NTV, Berkeley repeatedly argues for local versions of this thesis (that there is no common idea of distance (*NTV* §41), magnitude (*NTV* §79), or situation (*NTV* §93)) using variants of the question posed by William Molyneux to John Locke, namely, whether a congenitally blind person that suddenly acquired vision would be capable of identifying a cube from a sphere using only his new acquired sense. Drawing on this question and defending radically negative an-

* Article written as part of the research project *La pirámide visual: evolución de un instrumento conceptual*, supported by the Fondo de Investigaciones de la Universidad del Rosario (FIUR) in Bogotá, Colombia. I'm thankful to Carlos Cardona, Ignacio Ávila, Nicolás Montenegro, Susana Restrepo, and Juliana Gutiérrez for their comments on previous versions of this paper. All errors and obscurities remain mine.

[1] From now on, citations of Berkeley's *Essay* will be labeled *NTV*.

swers, Berkeley questions, not only the capacity of recognizing shapes, but also of recognizing distance, magnitude, situation, and any other idea whatsoever (*NTV* §128). I will call variations of the form "If a congenitally blind man were to recover vision, would he be able to identify X by vision alone?" (including the ones proposed by Berkeley) "Molyneux-type questions", and will reserve the label "Molyneux's question" to Molyneux's own formulations (I present two different formulations below), unless otherwise stated.

In reading Berkley's NTV, the reader might sometimes feel at odds with the argumentative role that Molyneux-type questions play throughout the text. At the end of NTV, Berkeley argues for a general Heterogeneity Thesis using three arguments, one of which is a negative answer to a generalized Molyneux-type question. This suggests that at least this last negative answer to a Molyneux-type question is logically prior to (i.e. is a premise of) the Heterogeneity Thesis. However, in motivating his arguments for local heterogeneity regarding distance, magnitude, or situation, Berkeley seems either to be presupposing Heterogeneity, or to be lacking support for his expected negative answers to Molyneux-type questions. In other words, it appears that the various negative answers to Molyneux-type questions that appear before the general Heterogeneity Thesis depend on such claim or are unwarranted. If this is so, then Berkeley's NTV runs the risk of circularity or is based on an undefended assumption.

In what follows, I will suggest a reading of Berkeley's NTV and Molyneux-type questions that saves these problems. I argue that the Heterogeneity Thesis works, not as a consequence of Berkeley's theory, but as its phenomenological basis. And if this is so, then the Molyneux-type cases explored in NTV can be read as thought experiments (*Gedankenexperimente*) illustrating, but not logically supporting, this basis. This implies that we should read Berkeley's negative answers to Molyneux-type questions as based on the Heterogeneity Thesis, and not the other way around.

I will first show the argumentative problems in Berkeley's NTV, and show how an elucidation of Molyneux-type questions becomes vital to escaping them. Second, I will present some remarks on thought experiments and apply them to Molyneux-type questions. Third, I will show a reading that, while putting the Heterogeneity Thesis at the basis, escapes circularity and offers a more solid interpretation of Berkeley's theory of vision.

2. The argumentative strategy in Berkeley's NTV

2.1. Berkeley's arguments in the NTV

Berkeley opens his NTV by presenting his objective:

> My design is to show the manner wherein we perceive by sight the distance, magnitude, and situation of objects. Also to consider the difference there is betwixt the ideas of sight and touch, and whether there be any idea common to both senses. (*NTV* §1)

Following this order, Berkeley intends to convince us that distance, magnitude, and situation, are not perceived by lines and angles alone, as he believes the tradition has supposed.[2]

[2] It is not true that the whole tradition before Berkeley considered lines and angles alone to be the manner in which we perceive distance by sight. For example, Descartes also considers the sensation of movement of the eyes as one of the contributing factors. See (Descartes 1637, AT VI 137).

For the time being, allow me to consider only the arguments for a local heterogeneity thesis regarding distance, although similar arguments can be run on Berkeley's discussion of magnitude and situation.

The first sections of NTV are then dedicated to Berkeley's attack on traditional optics. We can summarize his argument as follows:

(P1)	An idea is perceived either by itself or by means of some other idea.	(*NTV* §9)
(P2)	If an idea is not perceived itself, it cannot be the means of perceiving any other idea.	(*NTV* §10)
(P3)	Distance is not itself perceived by sight.	(*NTV* §11)
(C1)	By (P1) and (P3), distance must be perceived by means of some other idea.	
(P4)	The lines and angles supposed by the tradition are not themselves perceived.	(*NTV* §12)
(C2)	By (P2) and (P4), the lines and angles supposed by the tradition cannot be the means of perceiving any other idea.	
(C3)	By (C1) and (C2), distance cannot be perceived by means of the lines and angles supposed by the tradition.	

How, then, do we perceive distance by sight (if we do at all)? Berkeley's view is that it is by means of contingent correlations between the ideas of touch and the ideas of sight (cf. *NTV* §45). It is by associating certain visual cues to tangible ideas that we have an idea of distance. From this, Berkeley concludes:

> From what has been premised it is a manifest consequence that a man born blind, being made to see, would at first have no idea of distance by sight. The sun and stars, the remotest objects as well as the nearer, would all seem to be in his eye, or rather in his mind. The objects intromitted by sight would seem to him (as in truth they are) no other than a new set of thoughts or sensations, each whereof is as near to him as the perceptions of pain or pleasure, or the most inward passions of his soul. For our judging of objects perceived by sight to be at any distance, or without the mind, is entirely the effect of experience, which one in those circumstances could not yet have attained to. (*NTV* §41)

This is Berkeley's first negative answer to a Molyneux-type question: a man born blind, after recovering vision, could not recognize objects as being at a distance by sight alone. In other words, distance, at least in the sense of it being a line between objects and the observer, is not an idea that we perceive by sight. Berkeley then runs similar arguments with analogous conclusions regarding magnitude and situation. As with distance, magnitude and situation are ideas that we do not perceive by sight. If we were pressed to believe that they refer to external objects, they would be ideas perceived by touch.[3] Later, in the *Principles of Human Knowledge* (Berkeley 1710), he will reject that touch can give us such information at all.

[3] More precisely, Berkeley believes that we should distinguish between visual distance, magnitude, and situation, and their tangible counterparts. Each one of them is composed by what he calls *minima visibilia* (minimally perceivable visual points) and *minima tangibilia* (minimally perceivable tangible points), respectively. In these arguments, however, he is referring to tangible distance, the number of *minima tangibilia* between an object and an observer, instead of visible distance, the number of *minima visibilia* between two points in our visual field. See (*NTV* §54)

Finally, towards the end of the NTV, Berkeley argues for a more general separation between the ideas of sight and touch. He first rejects the notion of abstract ideas (*NTV* §§122-126), an attack that will be repeated in a stronger form in the *Principles of Human Knowledge*. Then, he writes:

> It having been shown that there are no abstract ideas of figure, and that it is impossible for us by any precision of thought to frame an idea of extension separate from all other visible and tangible qualities which shall be common both to sight and touch, the question now remaining is: whether the particular extensions, figures, and motions perceived by sight be of the same kind with the particular extensions, figures, and motions perceived by touch? In answer to which I shall venture to lay down the following proposition: *The extension, figures, and motions perceived by sight are specifically distinct from the ideas of touch called by the same names, nor there is any such thing as one idea or kind of idea common to both senses.* (*NTV* §127, emphasis in the original)

Berkeley offers three arguments for this proposition. The arguments appeal to (i) the newly sighted man's incapability to establish similarity relations between the ideas of sight and touch (*NTV* §128); (ii) the claim that light and colors are proper sensibles of vision, not perceivable by touch (*NTV* §129-130); and (iii) the impossibility of adding visual and tactile quantities (*NTV* §131). Only the first one should concern us here.

This first argument begins with the claim that when we classify two ideas under the same class, we do so in virtue of a similarity relation. Berkeley believes that Molyneux's patient would be incapable of establishing such similarities, and on the contrary, would believe that there is nothing in common between the ideas of touch and the ideas of sight (*NTV* §128). This, he thinks, is argument enough to conclude that there is no idea perceivable by two sensory modalities. Berkeley, however, offers no argument for his expectation that Molyneux's patient should answer negatively. Moreover, the only argument he can offer is based in turn on the Heterogeneity Thesis he is trying to defend with his argument. Let's examine this in closer detail.

2.2. The argumentative problem in NTV

As seen in NTV §127, Berkeley believes that the arguments mentioned above are only demonstrations to clarify what has been already laid out in the rest of the text. What we find in the rest of the text, however, are local versions of the Heterogeneity Thesis. As mentioned above, for example, there is an argument for the claim that distance is not an idea common to sight and touch. But this argument, as well as the others before §127, (i) only show that there exist *some* ideas that are not common to these sensory modalities, thus providing no support for a Strong Heterogeneity claim, and (ii) are based on either undefended negative answers to Molyneux-type questions, which puts the NTV on a weak foundation, or are based on Strong Heterogeneity, which then puts the NTV at risk of circularity.

Regarding the first issue, some have then distinguished between two different Heterogeneity Theses in the NTV (Muehlmann 2008; Wilson 1999). The first thesis is sometimes dubbed *Weak* or *Numerical Heterogeneity*, and only states that "we never feel and see one and the same object" (*NTV* §49). As Muehlmann (2008) points out, this thesis says nothing about the qualitative difference or sameness of vision and touch; it just states that there is at least one property that is not shared by the objects of vision and touch, such that they are numerically distinct. A second version, dubbed *Strong* or *Qualitative Heterogene-*

ity is the claim that there is no property in common to the ideas of vision and touch, and so they are different in quality.[4] Put this way, the problem becomes one of jumping from one thesis to the other. The local arguments for Heterogeneity serve as arguments for Weak Heterogeneity, but not for Strong Heterogeneity. Under this reading, the NTV's argumentative strategy bears on an unwarranted generalization from Weak Heterogeneity claims to Strong Heterogeneity.

I will argue strongly against this reading in the third part of the text. For the time being, I will point out that if Berkeley wants to defend a Strong Heterogeneity Thesis, as I understand he does, then his view in NTV would be that Weak Heterogeneity does not even make sense, and so reading Berkeley in the light of this distinction blurs the general claim he is making. But leaving this argument aside, there is still a problem regarding Berkeley's argumentative strategy, even if he intends to defend weaker Heterogeneity Theses in the first parts of NTV, namely, that either his negative answers to Molyneux-type questions are based on these heterogeneity claims, which in turn are not justified, or these negative answers support the heterogeneity claims, but then we have no support for these answers themselves.

Now, let us turn to the more pressing issue, the dilemma between circularity and an unwarranted foundation. Consider again, for instance, Berkeley's arguments for Weak Heterogeneity about distance, i.e. the claim that distance is not a common idea between vision and touch. As it is phrased by Berkeley, the negative answer to Molyneux's question about distance stems from "what has been premised" (*NTV* §41), which is that distance is not perceived by lines and angles, but by contingent associations. But why, if not by lines and angles, should distance be perceived by these associations? All Berkeley has shown is that lines and angles are not good candidates for the visual recognition of distance, but he does not have an argument for the claim that there is heterogeneity in this regard. He could, on the one hand, argue for a Heterogeneity claim if we grant that Molyneux's patient could not recognize distance visually, which would show that vision alone is not sufficient for said recognition. But in this case, his negative answer to this Molyneux-type question would need support, a support that it would lack at this point of the NTV. On the other hand, Heterogeneity claims (especially Strong Heterogeneity) might support a negative answer to the Molyneux-type case about distance, but then these claims would, first, need further defense that cannot come from negative answers to Molyneux-type cases, and second, would result in the NTV begging the question. Thus, either Berkeley's argument is based on unsupported negative answers to Molyneux-type cases, or it is based on Heterogeneity claims which would make the NTV circular.

This problem becomes dramatic at the end of the NTV, as it becomes generalizable to all cases. As Wilson (1999) claims, the first argument for the Strong Heterogeneity Thesis is based on an assumption that Molyneux's patient would always answer negatively to any Molyneux-type question. If Berkeley thinks that this has been already proven by what has been said in the rest of the NTV, as he says it has (*NTV* §127), then it is either proven by negative answers to local Molyneux-type cases, which in turn lack support and thus weakens the foundations of the theory, or it is proven by a set of Weak Heterogeneity Theses (presumably for every idea, thus being implicitly equivalent to Strong Heterogeneity) or Strong Heterogeneity itself, making the text circular.

---

[4] I am indebted to Ignacio Ávila for pointing out this distinction in Berkeley's NTV.

In what follows, I will argue that we can make sense of the presupposition of a Strong Heterogeneity Thesis without risking this kind of circularity. We can do this by reading these Molyneux-type questions, not as arguments derived from or in favor of the Heterogeneity Theses, but as thought experiments that illustrate them. To do so, let's first lay some ground to analyze thought experiments and their function in scientific theories.

3. Remarks on thought experiments

Before discussing how Molyneux-type questions can be treated as thought experiments and what type of experiments they are, we must gain some clarity on what thought experiments are. To do this, I intend to (i) distinguish between an instantiated experiment, an imagined experiment, and a thought experiment; (ii) consider some of the possible roles thought experiments play in scientific theories; and (iii) present a taxonomy of thought experiments that will allow us to classify Molyneux-type questions and determine their role in Berkeley's NTV.

Let's begin with the distinction between an instantiated, an imagined, and a thought experiment. At a first glance, a thought experiment is distinguished from an instantiated experiment insofar as the first one may not be empirically realized (either because of its instantiation being unnecessary or even impossible), whereas the second works precisely because of its being carried out. Cases of the first kind are, for instance, experiments where frictionless planes are assumed. These experiments cannot be empirically realized, as such planes do not exist, but they may still play some role in scientific reasoning. Cases of the second kind are most scientific experiments, where real situations are created and controlled and hypotheses are tested against experience. In a nutshell, a thought experiment works in the realm of scientific imagination, while an instantiated experiment works in the realm of experience or reality.

However, this definition falls short, since it would imply that a merely imagined experiment in the mind of a scientist would qualify as a thought experiment (e.g. a scientist planning to do an experiment next week). Some authors have then drawn a further distinction between imagined experiments and thought experiments (Brown 1986; Sorensen 1992). Put broadly, in an imagined experiment, the primitive (or basic) suppositions of the theory are already given; in a thought experiment, on the other hand, they are the matter under discussion.

Consider the case of digging a hole that goes through Earth and asking what would happen if we let an object fall through it: would it be suspended in the center of the Earth, or would it go all the way to the other side of the hole? This experiment, despite being impossible to carry out empirically, discusses no new conceptual scheme, no new basis for a scientific theory. It is simply appealing to imagination, but it is not proposing any new way of explaining, predicting, or even conceiving phenomena. In other words, it is no more than an *imagined experiment*.

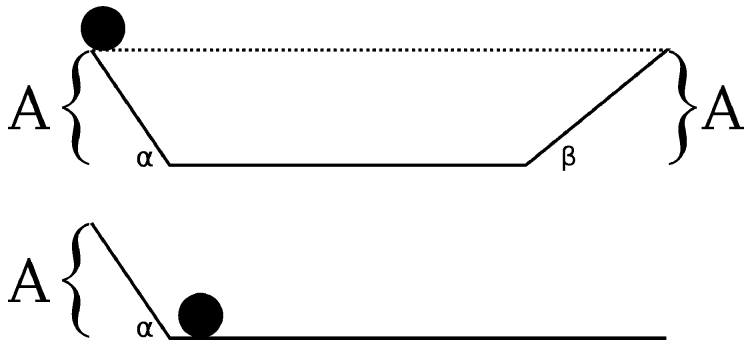

Figure 1

Illustration of Galileo's thought experiment for the principle of inertia.
Top: first scenario where the sphere rolls through the system and reaches height A.
Bottom: second scenario of the system, where $\beta = 0^{o}$.
The sphere presumably takes an infinite amount of time rolling up to height A,
and so movement would be its natural state until something stops it

Now consider a contrasting case: Galileo's thought experiment to argue for his principle of inertia (cf. Galilei 1632, 145-147). In this experiment, there is a basic supposition under discussion: that objects are to be considered as naturally tending towards movement (instead of the Aristotelian idea that they tend towards rest). Galileo asks us to imagine a frictionless system as seen in Figure 1. In the first scenario of this system (top), a sphere rolls down an inclined plane, then through a flat region, and finally rolls up another inclined plane. Presumably, the sphere must reach the same height it had at its initial state (assuming there is no energy loss), i.e. height A. If we reduce angle β slightly, the sphere will take longer to reach height A than it did before. Now, if we reduce angle β such that $\beta = 0^{o}$, when will the sphere reach height A?

Presumably, in this second scenario, the sphere would roll *ad infinitum*, due to the impossibility of it reaching height A. If Galileo succeeds in convincing us that this is the case, we would be forced to say that movement is the natural state of objects, or in other words, that an object will move until some force stops it. In contrast, the Aristotelian interpretation of this system would assume that the sphere should stop as soon as there is no force acting on it, making it continue its course. In this experiment, then, there are two colliding world pictures: the Galilean, in which movement is the natural state of things, and the Aristotelian, in which rest is the natural state. In this sense, a thought experiment such as this one, besides being a product of imagination and not being decided through experience, discusses the basic presuppositions of a scientific theory.

Now, having a distinction between instantiated, imagined, and thought experiments, we can ask: what is the function of thought experiments? The discussion on this topic is broad, and I do not intend to enter into the epistemological debates regarding these experiments. However, with the distinction itself we can hint an answer to the question: a thought experiment discusses the basic presuppositions of a theory. Some authors debate whether this produces a priori knowledge (Brown 1986) or introduce new research ques-

tions by modifying the ideals of natural order (Sorensen 1992). For the time being, it is not necessary to decide between one or the other. It suffices to say that in any of these cases, thought experiments are about basic presuppositions, that is, they sell a different world view, even if this is meant as a priori knowledge or a change in ideals of natural order.[5] This formulation, although vague, allows for a reconstruction of Molyneux-type questions as thought experiments. But before going into this, let's explore a taxonomy of thought experiments that will gain us some clarity as to the particular types of thought experiments there are, and the function of the one we are interested in.

For the purpose of my argument, I will use Brown's (1986) taxonomy, although with some modifications. This taxonomy covers a number of thought experiments, offers an explanation as to why thought experiments can yield a change of theory, and it can subsume other taxonomies such as Popper's (1935). Brown distinguishes—although the distinction is not exclusive—between «destructive» and «constructive» thought experiments. Destructive thought experiments, as their name suggests, destroy or contradict a theory (for example, Galileo's experiment described above works as a destructive experiment, as it intends to contradict Aristotelian physics; other examples include Maxwell's demon (1871) to show that the second law of thermodynamics can be violated). Constructive thought experiments, again following their name, settle the foundations of the theory (Einstein's (1956) elevator, Newton's (1687) bucket, and Schrödinger's cat (1935) are famous examples of such cases).

Among constructive thought experiments, Brown distinguishes between «conclusive», which work as full blown arguments in favor of a theory, and «conjectural», which only add plausibility to it (Brown 1986, 5). I will use Brown's terms in a somewhat deflated sense. I do not believe that thought experiments can be conclusive against a theory; as powerful as they may be, it takes more than a thought experiment to lead us to abandon our theoretical commitments. This distinction, however, seems to capture something else that will prove useful in what follows, namely, that conjectural thought experiments have different standards of justification than those of conclusive ones. Let me explain what I mean by this.

In conclusive experiments, it suffices that the possibility of a phenomenon is accepted without further justification. For example, Einstein's elevator thought experiment is based on the possibility of there being an elevator in which we cannot tell whether there is a gravitational field or acceleration, but it does not need to justify the possibility of such an elevator existing. In other words, Einstein does not need to convince us that it is possible for such an elevator to exist. It is a matter of sympathizing with the proposed situation or not sympathizing at all.

In contrast to conclusive experiments, conjectural experiments do require an argument to accept the phenomenon in question. In other words, the phenomenon itself does not suffice for the experiment to succeed, but we need to justify that the phenomenon can indeed happen. Consider Newton's bucket experiment, where we need to justify why water

[5] It is not relevant if this change in world view is progressive or not, if it produces new knowledge or not, etc. I would rather remain agnostic regarding this question, although I sympathize with the Kuhnian (Kuhn 1964) version of thought experiments, according to which they offer new conceptual schemes or paradigms.

in a bucket in a vacuum would indeed move the way Newton thinks it does. Here, Newton does need to convince us that the phenomenon in question (water moving in this and that way) is actually possible. Moreover, as Brown points out, a way of contesting this kind of experiment is by rejecting the possibility of such phenomenon (as did Berkeley (1710) and Mach (1883) with Newton's bucket experiment). If this is the case, then conjectural experiments require a different kind of justification than do conclusive experiments.

In the case of Molyneux-type questions, at least in Berkley's theory of vision, I will argue that they work as thought experiments and not merely as imagined ones. Furthermore, I will claim that their function is to add plausibility to the Strong Heterogeneity Thesis, and so I will classify them as constructive conjectural thought experiments.

4. Strong Heterogeneity and Molyneux-type questions

4.1. Are Molyneux-type questions imagined or thought experiments?

Let's now turn to Molyneux's Question. In its first version, Molyneux asked Locke not only if a man born blind that recovers vision would be able to identify a cube and a sphere using only his vision, but also whether he could say how far away they lie from him. In Molyneux's words:

> A Man, being born blind, and having a Globe and a Cube, nigh of the same bignes, Committed into his Hands, and being taught or Told, which is Called the Globe, and which the Cube, so as easily to distinguish them by his Touch or Feeling; Then both being taken from Him, and Laid on a Table, Let us suppose his Sight Restored to Him; Whether he Could, by his Sight, and before he touch them, know which is the Globe and which the Cube? Or Whether he Could know by his Sight, before he stretchd out his Hand, whether he Could not Reach them, tho they were Removed 20 or 1000 feet from him? (Molyneux 1688, quoted in Degenaar 1996, 17)

In a second version of the question—the one that would reach Locke's hands—Molyneux (1693, quoted in Degenaar 1996, 21) repeats only the first part (about identification of shape), omitting the second (recognition of distance) altogether. I will consider the first version, since it includes two different formulations and thus serves a more general examination. The same argument, however, can be applied to the second version, as well as the other Molyneux-type questions considered by Berkeley.

Molyneux's question asks us to imagine something that, in principle, is not impossible: a congenitally blind man that recovers vision and must identify a cube from a sphere or objects at a distance. At the time of formulation, there was no empirical realization of this experiment (cf. Degenaar 1996). This can give the impression that Molyneux's question, as well as Molyneux-type questions in general, were just imagined experiments waiting for an instantiation.

Nonetheless, it is worth noting that the situation envisaged by Molyneux-type questions is not that easy to conceive as it would seem at first. It is not only difficult for there to be a congenitally blind person recovering vision (Valvo (1971) estimated that less than twenty cases had been found in the last 1000 years, although the number has presumably increased in the last decades), but this person would not be able to use his vision immediately. Presumably, the first experiences such a patient would have after opening her eyes

would be of pain (consider the adaptation period when sighted people see light after being adapted to a dark environment). But even after this pain period (which would count as an adaptation to visual stimuli, thus invalidating the patient as an optimal candidate for the experiment), it is still debatable whether her visual faculties are sufficient for a robust identification of shape or distance. To the best of our current empirical evidence, the best cases we have to test Molyneux-type questions have been not completely blind from the beginning, but blinded at a young age; some have even been able to identify the direction of sources of light (which would imply using visual cues and mapping things in an egocentric space); and lastly, they have had some period of time to gain familiarity not only with visual experience, but with some of the tests they were presented (Cheselden 1727; Connolly 2013; Gallagher 2005; Held *et al.,* 2011; Huber *et al.,* 2015; Meltzoff, and Borton 1979; Sacks 1995; Schwenkler 2012).

Such is the case, for example, of the study by Held *et al.* (2011), where they tested five subjects, four of which presented dense congenital bilateral cataracts, and the other one presented bilateral congenital corneal opacities. All of them received surgery to give them back their visual faculties. Before their interventions, none of them were able to discriminate form. After the intervention, the subjects could not immediately identify tactile with visual shapes, even though they were successful after some time gaining familiarity with the stimuli. This lead the researchers to conclude that an answer to Molyneux's question was "likely negative" (Held *et al.,* 2011, 552). However, none of the subjects were blind from birth, and their relative success in identification tasks could be explained by recourse to familiarity instead of cross-modal transfer. Therefore, it seems that testing Molyneux's question empirically might involve important methodological constraints that preclude a clear experimental answer to the question.

Due to these methodological and empirical constraints, Jacomuzzi, Kobau, and Bruno (2003) argue that Molyneux-type questions in general—not only those present in the NTV—should be read as thought experiments. They consider a wide range of empirical evidence and possible reformulations of Molyneux-type questions, and press the idea that Molyneux's question is a thought experiment due to the impossibility of obtaining the situation envisaged by it. In their own words, it is "a form of mental exercise which is such by necessity, not because of technological contingencies but because the conditions of the hypothetical experiment will never be realized in practice" (Jacomuzzi, Kobau, and Bruno 2003, 264).

Even though I share a similar view, my proposal differs from theirs in two ways. First, I do not make any claims regarding Molyneux's question, but only regarding Molyneux-type questions in Berkeley's NTV. Second, and most importantly, the methodological and empirical constraints Jacomuzzi *et al.* consider are not sufficient to justify their interpretation of Molyneux's question. Jacomuzzi *et al.* argue that Molyneux's question is a thought experiment because there cannot ever be a realization in practice, but an alternative is still available, namely, that Molyneux's question is an imagined experiment with an "infinite delay", as Sorensen (1992, 193) claims. In this case, it would not be any different from the case of imagining digging a hole that goes through Earth and letting an object fall through it. Both would be imagined experiments that cannot ever be instantiated, but that do not say anything about the bases of scientific theories at all; if anything, they are mere food for thought. Jacomuzzi *et al.* reject this view by stressing that the instantiation is not obtainable in principle, but their argument does not distinguish between the case of Molyneux's

question and the hole in the Earth one. Thus, further reasons are needed to exclude this interpretation.

The motivation behind interpreting Molyneux-type questions in the NTV as imagined experiments has, of course, some plausibility. In *The Theory of Vision Vindicated and Explained* (1732), Berkeley briefly alludes to Cheselden's (1727) first experimental report on a Molyneux-type case in which the patient did not pass Molyneux's test, that is, he could not identify shapes or distances using vision alone. He states that "[...] by fact and experiment, those points of the theory which seem the most remote from common apprehension were not a little confirmed, many years after I had been led into the discovery of them by reasoning" (Berkeley 1732, 304). This might lead to the idea that Berkeley intended his Molyneux-type cases to be delayed experiments that were later confirmed empirically.

This interpretation, however, can still be resisted. Berkeley holds he has already discovered the negative answer to Molyneux-type questions by reasoning alone. If this is so, then he did not think he would need empirical confirmation for what he thinks he already found out. This is consistent with the fact that only until the end of *The Theory of Vision Vindicated and Explained* does he comment on Cheselden's work, and he only does so in a paragraph. Instead, his efforts lie in convincing us of his expectations by using a priori arguments.

Furthermore, in the NTV, nowhere are Molyneux-type questions invoked as experiments waiting for an empirical instantiation. For Berkeley, all of these cases can be solved by reasoning. He even urges us to put ourselves in the shoes of a Molyneux patient and to "frame true conceptions of what might reasonably be supposed to pass in his mind" (*NTV* §92). Thus, his recourse to Molyneux-type questions in the NTV is not that of an experimental psychologist or something of the sort, but of a philosopher arguing from a priori grounds.

Lastly, it is plausible that Berkeley would have rejected a positive outcome from an experimental instantiation of Molyneux-type questions. Even if Molyneux's patient reports being able to identify visual shapes, distances, magnitudes, etc., with their alleged tangible counterparts, there are reasons to suppose that Berkeley would have been skeptical. Not only is his take on Molyneux-type cases *a priori*, but he takes them to say something about what Molyneux's patient would mean by the terms «shape», «distance», «magnitude», etc. Consider, for instance, the following passage in the *Philosophical commentaries*:

> By extension, a blind man would mean either the perception caused in his touch by something he calls extended, or else the power of raising that perception, [which] power is without in the thing term'd extended. Now he could not know either of these to be things visible till he had try'd. (Berkeley 1707, §100).

Berkeley's view here is that for a Molyneux patient, the meaning of terms such as «extension» would have to be fixed by his tactile sensations, such that he cannot even say meaningfully that he sees and touches the same thing (i.e. that what he sees is "extended" at all). If this is the case, even if there was a congenitally blind person that recovered vision and reported to identify visual shapes with tangible shapes (and other visual phenomena with tangible phenomena), Berkeley would argue that his/her reports are meaningless. Thus, experimental instantiations would not say anything relevant to Berkeley's theory of vision; Cheselden's report is just a lucky coincidence that confirmed his view, but, as Berkeley would take it, was not necessary for it.

If I am right, then we can resist reading Molyneux-type questions in the NTV as imagined experiments and proceed to consider them as thought experiments. In the previous section, I deployed a non-exclusive distinction between constructive and destructive thought experiments. In the case of the NTV, Molyneux-type questions can be classified in either way. On the one hand, it intends to contradict the doctrine of common sensibles defended by the tradition (specifically by Locke (1689)). If Berkeley is right in his negative answers to Molyneux-type cases, it follows that there cannot be any idea common to two sensory modalities. The power of this cases, however, does not stop at rejecting the tradition of common sensibles, but, I will argue, it intends to illustrate the basis of Berkeley's theory, Strong Heterogeneity.

4.2. Strong Heterogeneity as the basis of the NTV

Before defending the claim that Molyneux-type questions are constructive thought experiments that illustrate Strong Heterogeneity, we should accept that this thesis is one of the basic presuppositions of the whole theory, instead of a consequence of it. To argue for this claim, let's examine a reconstruction of Berkeley's theory.

In *The Theory of Vision Vindicated and Explained*, Berkeley claims that in the NTV he has followed a method by which "from false and popular suppositions, men do often arrive at truth" (Berkeley 1732, §38), i.e. a method of analysis. He then sets out to reconstruct his theory by a method of synthesis, a method of "delivering science or truth already found [in which] the conclusions of analysis [are] assumed as principles" (ibid.). In this new reconstruction, he takes the main principle to be the idea that vision is the language of the Author of Nature, that is, that vision is but a mere sign to anticipate tactile ideas, a sign configured by God in order to aid survival (*NTV* §147). From this principle, it follows that there is no idea in common to both vision and touch, that is, Strong Heterogeneity. And from Strong Heterogeneity, the rest of the arguments follow, or so Berkeley thinks.

If we read the NTV in the light of this new order, then the distinction between two different heterogeneity claims dissolves. The distinction works if we can read Berkeley as first defending the claim that there is a quantitative or numerical difference between the objects of vision and touch, and only afterwards defending a stronger qualitative version of that claim. But quantitative or numerical differences only make sense if we are comparing two qualitatively similar objects. In the words of Russell:

> To know whether, or by what means, comparison is possible, we must know the qualities of the things compared and of the medium in which comparison is effected while to know that quantitative comparison is possible, we must know that there is a qualitative identity between the things compared, which again involves a previous qualitative knowledge. (Russell 1897, §61)

If we take Russell's consideration seriously, we have to say that, if Berkeley is holding a Strong Heterogeneity claim, it does not even make sense to hold Weak Heterogeneity claims altogether, as the means of comparison would not even be available. Thus, there is only a Strong Heterogeneity thesis that is analyzed all throughout the NTV, instead of different theses.

My reading of Berkeley's NTV would then be as follows. First, there are basic assumptions regarding the nature of vision (the claim that vision is the language of the Author

of Nature) and from there a phenomenological principle to study perception, namely, Strong Heterogeneity. I say «phenomenological» since Berkeley's whole analysis of the way in which we perceive is guided by Strong Heterogeneity: he analyzes each sensory modality independently of the others, and only after this analysis he inquiries about the relation between them. Molyneux-type cases, I will argue, offer an illustration of how to carry out a phenomenology of perception. They provide an imaginary case (i.e. a thought experiment) that makes it clear how a sensory modality should be considered independently of the others. And, if my previous argument is correct, we cannot read Berkeley as carrying out this analysis locally, that is, as defending local heterogeneities. It follows from this that we cannot read these Molyneux-type case questions as instances of Weak Heterogeneity, but that we should read these cases instead as illustrations of what Strong Heterogeneity amounts to.

4.3. Molyneux-type case questions in the NTV as constructive conjectural thought experiments

So far, I have argued that (1) Molyneux-type case questions in the NTV are not imagined experiments, but thought experiments, and (2) that Strong Heterogeneity is at the base of his theory of vision (as a corollary, I argued that there is only Strong Heterogeneity in the NTV). If these arguments are correct, then the next question is to identify what kind of thought experiments Molyneux-type case questions are.

Prima facie, they are thought experiments against the claim that there can be an idea common to two or more sensory modalities, i.e. the tradition of common sensibles. This means that, for one, they are destructive thought experiments. However, there is more to the NTV than just abandoning the tradition of common sensibles; there is a proposal regarding how to conceive perception in general and vision in particular, namely, by conceiving every sensory modality as independent from the others, i.e. Heterogeneity. How does Berkeley push forward his Heterogeneity claim? As it should be clear by now, he invokes Molyneux-type cases that intend to show what it means to conceive sensory modalities as different from each other. In his view, any reader that accepts his take on Molyneux-type cases should accept Strong Heterogeneity. If this is so, then these Molyneux-type cases work as *constructive* thought experiments in support of the Strong Heterogeneity claim.

What kind of constructive thought experiments are these? It is important to remember Berkeley constantly argues for the possibility of these cases. As I stated above, Molyneux-type cases are not as easy to devise as one would initially believe, and they may even be impossible empirically. Berkeley then needs to provide a justification regarding why we should sympathize with his negative answers to Molyneux-type cases. According to the previous taxonomy, it is a conjectural thought experiment that requires such justification (in contrast to conclusive ones). Berkeley's negative answers to Molyneux-type questions are not a matter of mere sympathizing with his view or not. The case is more analogous to that of Newton's bucket, in which there is room for debate regarding the support of the thought experiment itself. In the case of Molyneux-type questions, this is supported by the fact that several philosophers at the time rejected Berkeley's view by denying that Molyneux's patient would answer negatively to all the relevant questions (cf. Degenaar 1996). Thus, Molyneux-type questions in the NTV are constructive thought experiments of the conjectural type.

By taking Molyneux-type questions in the sense of conjectural constructive thought experiments, the argumentative problem dissolves. The problem that the traditional reading underlined was that the negative answers to Molyneux-type cases were either assumed without grounds, or supported by Strong Heterogeneity. In the first case, the NTV is based on a series of undefended assumptions; in the second, it is circular. I have now proposed to take Strong Heterogeneity as the basis of the whole theory. Specifically, I have claimed that Strong Heterogeneity acts as a general phenomenological principle that states that sensory modalities should be conceived of as separate from each other.

I have argued that this Strong Heterogeneity claim lies at the base, that is, is not a consequence, of Berkeley's theory. If this is so, then Molyneux-type questions are either logically posterior or equivalent to Strong Heterogeneity. In the first case, we read Strong Heterogeneity as a premise in an argument concluding with negative answers to Molyneux-type questions. This reading, however, does not account for the fact that Berkeley thinks that his Molyneux-type cases somehow support Strong Heterogeneity, and not the other way around. By reading Molyneux-type cases as arguments derived logically from Strong Heterogeneity, we miss Berkeley's strategy.

Thus, I propose to take the second case, that in which Molyneux-type cases are somehow equivalent to Strong Heterogeneity. If Molyneux-type cases and Strong Heterogeneity are equivalent, then the difference lies in the mode of presentation. Strong Heterogeneity is presented in the NTV as the general phenomenological principle, whereas Molyneux-type cases are presented as thought experiments related to that phenomenological principle. He takes them to illustrate what is meant by this Heterogeneity claim. These cases show situations in which it is at least plausible that a man born blind does not know how to apply tangible ideas to his newly acquired vision, and so they provide plausible cases of Heterogeneity. In other words, they show what a phenomenology based on Heterogeneity amounts to. Thus, by invoking these cases in a variety of situations, Berkeley is persuading us to buy Strong Heterogeneity. This, in contrast to the first reading, is consistent with Berkeley's belief that Molyneux-type cases somehow support Heterogeneity; they support it insofar as they illustrate it, as they clarify what the claim is.

If this is the case, then the NTV is not circular, insofar as the relation between Heterogeneity and Molyneux-type cases is not an argumentative one, nor is it based on negative answers to Molyneux-type questions simpliciter, but on these answers as illustrations of Heterogeneity. One possible objection is that this reading does not dissolve the argumentative problem in the NTV, but only that it results in the NTV being based on an unwarranted assumption of Heterogeneity. To this objection, I respond that while it is true that Strong Heterogeneity would lack justification, this is not problematic for the NTV. It is a commonplace in scientific theories to have some basic presupposition at their foundation, a presupposition that is ultimately a way of seeing the world. Just as one cannot offer arguments for a different world view, but only persuade others to share it, the same goes for the bases of scientific theories. This is precisely why thought experiments are often used to illustrate or support the bases of scientific theories; they are not just arguments (if they are at all), but, as I claimed above, they offer a different world view. In the case of the NTV, it is Strong Heterogeneity what is being sold as an alternative to our general views on perception, but it is not problematic to have it as a basis of the theory, as Strong Heterogeneity itself is a general principle.

To sum up, I have argued that Molyneux-type cases in the NTV are, not only destructive thought experiments against the tradition of common sensibles, but also constructive

conjectural thought experiments that illustrate what Strong Heterogeneity amounts to. This is supported by the claim that the NTV is not only concerned with rejecting common sensibles, but with providing a theory of vision based on a phenomenology of Heterogeneity. Berkeley's strategy to sell this theory is by using Molyneux-type cases to illustrate his basic claim, and thus Molyneux-type cases become constructive thought experiments. Furthermore, they are thought experiments in which there is room for debate regarding whether Molyneux's patient would indeed answer negatively, and so they are thought experiments of the conjectural kind. Lastly, I claimed that my reading of Molyneux-type cases as constructive conjectural thought experiments dissolves the argumentative problems shown in the first section, as it avoids reading the NTV as circular or as based on negative answers to Molyneux-type cases simpliciter, and instead proposes to read the NTV as based on Strong Heterogeneity and Molyneux-type cases as illustrations of it.

5. Conclusion

In this article, I have argued that Molyneux-type questions in Berkeley's NTV should be read as conjectural thought experiments illustrating his Strong Heterogeneity thesis, instead of as logical support for such a claim. I presented a reading of Berkeley's NTV that has been accepted by the tradition and that yields a problematic argumentative structure in the NTV. I then deployed a notion of thought experiment in distinction to instantiated and imagined experiments, and argued that there is a sense in which Molyneux-type questions can be seen as cases of the first class. Furthermore, I considered the interpretation defended by Jacomuzzi, Kobau, and Bruno (2003), and presented a different proposal insofar as their argument (i) relates to Molyneux's original questions instead of Berkeley's Molyneux-type cases, and (ii) is based on the mere claim that Molyneux-type cases are not empirically possible, which does not allow a distinction between imagined and thought experiments. In this sense, I not only refer to Berkeley's Molyneux-type cases, but I add a stronger argument than the one Jacomuzzi *et al.* provide.

Finally, I defended an alternative reconstruction of the NTV that allows Molyneux-type questions to play an illustrative role, and, by dissolving the distinction between a Strong Heterogeneity Thesis and Weak Heterogeneity Theses, allows for a more consistent reading of the whole theory. This reading, I claim, would save Berkeley's NTV from the argumentative problems presented in the first part, and provides a more solid reading of the theory. In seeing Strong Heterogeneity as one of the bases of the theory, there is no problem in accepting Molyneux-type questions as illustrations instead of arguments lacking proper support.

As an invitation for further research, I believe this reading not only to be of exegetical value. Molyneux-type questions have not only been considered widely by Berkeley, but they have played an important role in modern and contemporary philosophy of perception, and more recently in the fields of psychology and neuroscience. An analysis of these questions may allow us to distinguish different ways of phrasing them in such ways that we can identify which questions involved are of a philosophical or conceptual nature (as the defense of a Strong Heterogeneity Thesis), and which allow for empirical research to jump in. In my view, this is a distinction that is worth tracing, in order to avoid confusing empirical and theoretical questions, and to guide philosophical and scientific inquiry in better oriented paths.

REFERENCES

Berkeley, George. 1707. Philosophical commentaries. In *The Works of George Berkeley Bishop of Cloyne,* ed. Arthur Aston Luce and Thomas Edmund Jessop, Vol. 1, 1-140. Nendeln: Kraus Reprint, 1979.

Berkeley, George. 1709. An essay towards a new theory of vision (*NTV)*. In *Philosophical writings*, ed. Desmond Clarke, 1-66. Cambridge: Cambridge University Press, 2008.

Berkeley, George. 1710. The Principles of Human Knowledge. In *Philosophical writings*, ed. Desmond Clarke, 67-150. New York: Cambridge University Press, 2008.

Berkeley, George. 1732. The Theory of Vision Vindicated and Explained. In *Philosophical works*, ed. Michael Richard Ayers, 279-304. London: Tuttle Publishing, 1993.

Brown, James Robert. 1986. Thought experiments since the Scientific Revolution. *International Studies in the Philosophy of Science* 1, nº. 1: 1-15.

Cheselden, William. 1727. An Account of Some Observations Made by a Young Gentleman, Who Was Born Blind, or Lost His Sight so Early, That He Had no Remembrance of Ever Having Seen, and Was Couch d between 13 and 14 Years of Age. By Mr. Will. Chesselden, F. R. S. Surgeon to Her Maj. *Philosophical Transactions of the Royal Society of London* 35: 447-450.

Connolly, Kevin. 2013. How to test Molyneux's question empirically. *I-Perception* 4: 508-510.

Degenaar, Marjolein. 1996. *Molyneux's Problem: Three Centuries of Discussion on the Perception of Forms.* Trans. Michael J. Collins. Dordrecht, Netherlands: Kluwer Academic Publishers.

Descartes, René. 1637. Optics. In *Philosophical writings*, Trans. J. Cottingham, R. Stoothhoff, and D. Murdoch, vol. 1, 152-176. New York: Cambridge University Press, 1985.

Descartes, René. 1983. *Oeuvres de Descartes,* ed. Charles Adam and Paul Tannery. Paris: Librairie Philosophique J. Vrin, 1983.

Einstein, Albert. 1956. *Über die spezielle und die allgemeine Relativitätstheorie*. Berlin/Heidelberg: Springer-Verlag, 2009.

Galilei, Galileo. 1632. *Dialogue Concerning the Two Chief World Systems—Ptolemaic & Copernican*, Trans. Stillman Drake. 2nd ed. Berkeley: University of California Press, 1967.

Gallagher, Shaun. 2005. Neurons and neonates: Reflections on the Molyneux problem. Chap. 7 in *How the body shapes the mind,* 153-172. New York: Oxford University Press.

Held, Richard, Yuri Ostrovsky, Beatrice de Gelder, Tapan Gandhi, Suma Ganesh, Umang Mathur, and Pawan Sinha. 2011. The newly sighted fail to match seen with felt. *Nature neuroscience* 14: 551-553.

Huber, Elizabeth, Jason M. Webster, Alyssa A. Brewer, Donald I. A. MacLeod, Brian A. Wandell, Geoffrey M. Boynton, Alex R. Wade, and Ione Fine. 2015. A Lack of Experience-Dependent Plasticity After More Than a Decade of Recovered Sight. *Psychological Science* 26, nº. 4: 393-401.

Jacomuzzi, Alessandra C., Pietro Kobau, and Nicola Bruno. 2003. Molyneux's question redux. *Phenomenology and the Cognitive Sciences* 2: 255-280.

Kuhn, Thomas Samuel. 1964. A Function for Thought Experiments. In *The essential tension: Selected studies in scientific tradition and change,* 240-265. Chicago: The University of Chicago Press, 1977.

Locke, John. 1689. *An essay concerning human understanding*. Raleigh: Hayes Barton Press, 2005.

Mach, Ernst. 1883. *The Science of Mechanics,* Trans. Thomas Joseph McCormack. London: The Open Court Publishing Co, 1919.

Maxwell, James Clerk. 1871. *Theory of Heat.* New York: Dover Publications, 2001.

Meltzoff, Andrew N., and Richard W. Borton. 1979. Intermodal matching by human neonates. *Nature* 282, nº. 5737: 403-404.

Molyneux, William. 1688. A Problem Proposed to the Author of the Essai Philosophique Concernant L'Entendement. In *Molyneux's Problem: Three Centuries of Discussion on the Perception of Forms*, ed. Marjolein Degenaar, 17. Dordrecht: Kluwer Academic Publishers, 1996.

Molyneux, William. 1693. Dedication to the Royal Society. In *Molyneux's Problem: Three Centuries of Discussion on the Perception of Forms*, ed. Marjolein Degenaar, 21. Dordrecht: Kluwer Academic Publishers, 1996.

Muehlmann, Robert. 2008. Strong and Weak Heterogeneity in Berkeley's New Theory of Vision. In *New interpretations of Berkeley's thought*, ed. Stephen H. Daniel, 121-144. New York: Humanity Books.

Newton, Isaac. 1687. *Newton's Principia: The Mathematical Principles of Natural Philosophy,* Trans. Andrew Motte. New York: Daniel Adee, 1846.

Popper, Karl Raimund. 1935. *The logic of scientific discovery*. New York: Routledge, 2005.

Russell, Bertrand. 1897. *An essay on the foundations of geometry*. Cambridge: Cambridge University Press.

Sacks, Oliver. 1995. To See and not to See. In *An anthropologist on mars,* 102-144. London: Picador.

Schrödinger, Erwin. 1935. Die Gegenwärtige Situation in Der Quantenmechanik. *Naturwissenschaften* 23, nº. 48 (November 1935): 807-812.

Schwenkler, John. 2012. On the matching of seen and felt shape by newly sighted subjects. *I-Perception* 3: 186-189.

Sorensen, Roy. 1992. *Thought experiments*. New York: Oxford University Press.

Valvo, Alberto. 1971. *Sight Restoration after Long-term Blindness: The Problems and Behavior Patterns of Visual Rehabilitation.* New York: American Foundation for the Blind.

Wilson, Margaret Dauler. 1999. The Issue of "Common Sensibles" in Berkeley's New Theory of Vision. In *Ideas and mechanism: Essays on early modern philosophy,* 257-275. Princeton, NJ: Princeton University Press.

Juan R. Loaiza is a doctoral candidate at the Institut für Philosophie and the Berlin School of Mind and Brain at Humboldt-Universität zu Berlin. His research focuses in the philosophy of mind and cognitive sciences, particularly in the philosophy of emotions and the philosophy of perception. He previously worked as a Research Assistant at the Escuela de Ciencias Humanas at Universidad del Rosario (Bogotá, Colombia).

Address: Institut für Philosophie and Berlin School of Mind and Brain, Humboldt-Universität zu Berlin. Luisenstraße 56, 10099. Berlin, Germany. E-mail: loaiza.juan@hu-berlin.de