

可想象性论证及其问题^①

——评查尔莫斯的认知二维语义学

魏屹东 陈敬坤^②

(山西大学 科学技术哲学研究中心/哲学社会学学院 山西 太原 030006)

[摘要] 可想象性论证是心灵哲学中的一个重要问题。传统的可想象论证集中于物理主义与反物理主义之争，但没有达成一致意见。查尔莫斯运用二维语义学对传统的可想象性论证进行了修正，重建了可想象性与可能性之间的联系。我们发现查尔莫斯并没有真正解决这个问题，他的论证仍然有缺陷。

[关键词] 查尔莫斯 可想象性论证 CP 论题 二维语义学

[中图分类号] N031 [文献标识码] A [文章编码]

在当代心灵哲学中，关于可想象性的物理主义与反物理主义论证是一个重要问题。反物理主义论证（知识论证、解释鸿沟论证、可想象性论证、属性二元论论证）大多是从物理事实与现象事实之间存在认知鸿沟这一前提出发，得出二者之间存在本体论鸿沟的结论。物理主义者则坚持认为，认识论的前提并不能导出本体论的结论。查尔莫斯（Chalmers, D.J.）运用二维语义学对传统的可想象性论证进行了修正，重建了可想象性与可能性之间的联系，认为从认识论前提合理地推出本体论结论是可行的，从而对物理主义提出了强有力的挑战。通过深入分析查尔莫斯的可想象论证，我们发现这种基于二维语义的可想象性论证仍然存在明显的缺陷，并不能真正驳倒物理主义。

一、传统可想象性论证的困境

物理主义主张，关于世界的物理事实P和现象事实Q之间存在着必然蕴含关系，这意味着：如果两个世界在物理上同一，则它们在一切方面同一，不可能只有物理属性而没有相应的感受性（qualia）。物理主义者常常用克里普克的观点为他们的主张辩护。克里普克认为，先天性和必然性之间那种康德式的联系并不存在，陈述的必然真并不需要先天真作为保证。克里普克令人信服地论证了后天必然命题的存在，比如“水是H₂O”、“热是分子运动”这些固定指示词之间的同一就是后天必然的。物理主义者将这些后天必然性的例子与心身关系进行类比，认为P和Q之间的蕴含关系即使不是先天必然的，至少也是后天必然的^③，意识概念虽然不同于任何物理的或功能的概念，但或许我们的经验会发现二者的指称其实是一样的，正如水和H₂O、热和分子运动这些具有同一指称的名称一样。

物理主义的这一观点和我们的直觉相悖，因为具有物理属性而没有相应的感受性似乎总是很容易想象的，传统的可想象性论证正是基于这样一种直觉，

^① 教育部人文社会科学重点研究基地重大攻关项目“语境实在论”（07JJD20051）和山西省教育厅课题“语境实在论与认知科学的发展”（20083002）及山西省留学项目“科学认知的机制与表征模型研究（0905502）”成果之一。

^② 魏屹东（1958—），男，山西永济人，哲学博士，山西大学科学技术哲学研究中心/哲学社会学学院教授，博士生导师，近年来主要从事认知科学哲学研究。陈敬坤（1981—），男，湖北武汉人，山西大学哲学社会学学院在读博士，主要从事认知哲学研究。

^③ 这种观点一般称之为后天物理主义，查尔默斯区分了A、B、C三种类型的物理主义，他的反物理主义论证主要针对的是辩护能力最强的B类物理主义，也就是后天物理主义，后文中的物理主义如果不加说明都是指这种后天物理主义。

其基本思路是：

既然物理主义坚持 $P \rightarrow Q$ (P 蕴含 Q) 是必然的，这就意味着 $P \& \sim Q$ (仅具有物理属性而没有感受性) 形而上不可能，如果能证明 $P \& \sim Q$ 形而上可能，那么物理主义的观点显然就是错误的。由于“仅具有物理属性而没有感受性”的人在现实世界中是不存在的，因而是怪人 (zombie)，但我们却可以想象它们的存在。如果能够从怪人的可想象性推出其形而上可能性，就能顺利驳倒物理主义，因而，这种可想象论证又称为怪人论证，其论证形式如下：

- (1) 怪人可想象；
- (2) 如果怪人可想象，则怪人形而上可能；
- (3) 如果怪人形而上可能，则物理主义错误 (结论)。

这一论证有两个关键问题：一是怪人何以可想象；二是可想象的怪人何以可能。许多哲学家乐于承认怪人在某种意义上确实是可想象的，不过这里的“某种意义”显然太宽泛、太模糊。更重要的是，这一论证还依赖于一条形而上学的假设：可想象的即可能的，或者说可想象性蕴含可能性 (我们将这一假设称为 CP 论题)。CP 论题长期以来几乎被当作公理直接拿来运用，休谟就曾明确指出：“形而上学中有一条确立的公理，那就是：凡心灵能够清楚地想象的任何东西，都包含可能存在的观念，换句话说，凡我们所想象到的东西都不是绝对不可能的。我们能够形成一座黄金色的山的观念，由此就可断言，这样的一座山可能真正存在。我们不能形成一座没有山谷的山的观念，因此就认为这样的山是不可能的。”^{[1]45}

的确，当我们被问及某种事物是否可能时，我们总是试图去想象它，如果能够想象则认为是可能的，反之则不可能。但这一直觉显然并不像“两点之间直线最短”那样令人信服，因为可想象性是一个认知概念，而可能性属于模态概念，前提 (2) 存在着从认知领域到模态领域的跨越。可想象性的门槛越低，接受前提 (1) 就越容易，而接受前提 (2) 就越难。因此，对可想象性论证的“常规诘难” (standard objection) 就是反驳 CP 论题^{[2]15-23}。

对 CP 论题最具影响力的批评来自克里普克和普特南，他们在意义理论方面的工作直接切断了可想象性和可能性之间的联系。克里普克反对物理主义者将“水/ H_2O ”的同一性和心身关系进行类比，但他严格区分了认知领域和模态领域，切断了二者之间的联系。可想象性是一个认知概念，对于某主体是否可想象取决于他知道什么相信什么，或者是他使用了什么样的概念或表征方式。按照这种认知的理解，可想象性并不直接蕴含可能性，因为这时候可想象性取决于我们所能提供的概念和背景信息，一个命题即使在现实中是不可能的，我们也可以想象它。换言之，认知想象中可能的未必形而上可能。比如，我们说哥德巴赫猜想无论正确与否都是可想象的，但这两种情况只有一种是可能的，而“水不是 H_2O ”虽然可想象，但是不可能。普特南也指出，虽然就某种特殊的意义而言，“水不是 H_2O ”是可想象的，但逻辑上是不可能的，可想象性并不是可能性的依据^{[3]55-57}。

显然，如果 CP 论题不成立，那么可想象论证就立刻失去了支撑。支持可想象性论证必须面对这一严重的困难。查尔莫斯发现卡普兰、斯道纳克、埃文斯等人主张的二维分析方法对解决这一问题非常有帮助。他们认为索引词甚至名称和句子在不同的可能世界有不同的指称，并且不同程度地和认知意义联系在一起，这就意味着通过二维方法重建认知领域和模态领域之间的联系似乎是可行的。但是查尔莫斯认为他们大多把可能世界理解为表达的语境，而这种基于语境的分析非常有局限性，因为像“我在思维”、“语言存在”这样的表达式在任何语境中都为真，基于语境的解释并不能对这些句子的认知状态给予完全的

说明，也不能在表达和认知意义之间建立一种普遍的联系^{[4]71-75}。因此，查尔莫斯主张一种认知的理解，提出了一种认知的二维语义学解决方案。

二、认知二维语义学解决方案

一般认为，与一个表达式相关的首先是其外延。在二维语义学中，表达式的外延以两种不同的方式依赖于外在世界可能的性质或状态：其一，表达式的外延依赖于表达所处的现实世界的特征，即现实世界呈现的方式；其二，当现实世界的特征已经确定，表达式的外延还依赖于世界的虚拟特征。表达式与外延之间的这种联系是由表达式的内涵实现的，可以说，内涵就是从世界到外延的函项。

对应于这两种依赖方式，就有两种内涵和对可能性的两种不同理解：一是认知可能性，即把可能性理解为现实世界呈现的方式，每一种呈现方式是现实世界的一种可能的状态，或者说，是一种被视为现实的可能世界，这实际上是我们对现实世界可能发生的事实的认知可能假设，查尔莫斯把这些不能够先天排除的具体假设称为“情形”（scenario）。从情形到外延的函项就是第一内涵，又称为认知内涵；二是我们通常所熟悉的形而上可能性，即当现实世界已经固定的时候，我们把可能性理解为反事实的。从可能世界到外延的函项即第二内涵，又称为虚拟内涵。

基于这样一种二维语义，查尔莫斯对克里普克的后天必然性进行了分析。克里普克认为“水”这样的名称是固定指示词，如果水指称 H_2O 的世界（比如我们的世界）成为现实的，那么，水在所有可能世界指称 H_2O 。因此“水不是 H_2O ”是形而上不可能的。然而，尽管水不是 H_2O 形而上不可能，但认知上是可能的，因为我们不能先天排除一些认知上可能的情形，如孪生地球，在这个情形中，江河湖海中的液体物质是 XYZ。如果这个情形果真实现，那么，我们就必须承认“水不是 H_2O ”是真的，因此可以说，“水不是 H_2O ”的第一内涵在这个情形为真。这里关键在于区分形而上模态和认知模态，或者说，区分一般的虚拟赋值和对句子的认知赋值。第二内涵基于前者，第一内涵基于后者。所以，“水是 H_2O ”是后天必然的，因为它的第二内涵在所有可能世界为真（包括孪生地球），第一内涵在某些情形（如孪生地球）中为假。

可见，如果句子 S 是先天的，则 S 的第一内涵在所有情形为真。如果句子 S 不是先天的，则 $\sim S$ 在认知上是可能的。或者说，S 的第一内涵在某情形中真值为假。因此，S 的第一内涵在所有情形为真，当且仅当 S 是先天的，反之亦然。可以看到，在这里，第一内涵完全是以认知术语来描述的，查尔莫斯用这种方式对一般模态概念进行了定义：

（1）句子 S 是形而上必然的，当且仅当 S 的第二内涵在所有世界为真。

（2）句子 S 是后天必然的，当且仅当 S 的第二内涵在所有世界为真而第一内涵在有些情形为假。

（3）句子 S 是先天偶然的，当且仅当 S 的第一内涵在所有情形为真，第二内涵在有些世界为假。

（4）句子 S 是先天的（认知上必然的），当且仅当 S 的第一内涵在所有情形为真。

（5）“ $A \equiv B$ ”是形而上必然的，当且仅当 A 和 B 的第二内涵相同。

（6）“ $A \equiv B$ ”是先天的（认知上必然的），当且仅当 A 和 B 的第一内涵相同。

（1）和（4）定义了两种不同的必然性，其中（1）描述的实际上就是克里普克的观点，即，如果 S 在所有可能世界为真，则 S 是形而上必然的。这是对形而上必然性的传统理解。（4）则是（1）在认知领域的反映，它是查尔莫斯认

知二维语义的核心观点，因为在其它二维框架中，先天性和第一内涵（即认知内涵，在其他的二维框架中可能有不同的名称，比如，杰克逊将 *intension* 称为“A 内涵”）之间的联系是受约束的，而查尔莫斯使这种联系具有普遍性。在查尔莫斯看来，二维语义的核心立场就在于，如果涉及专名和自然种类词的后天同一，其第一内涵必定在某些情形中为假。

在（5）和（6）中，A 和 B 是任意同类型的表达，“ $A \equiv B$ ”是真句子当且仅当 A 和 B 有同一指称。如果 A 和 B 是单词，则 $A \equiv B$ 相当于“ $A = B$ ”；如果 A 和 B 是句子，则 $A \equiv B$ 相当于“A 当且仅当 B”。（6）很容易让人想起弗雷格的观点：“ $A = B$ ”不具有认知意义，当且仅当 A 和 B 含义相同。显然，与弗雷格一样，查尔莫斯也主张，在意义和认知领域之间存在某种联系。

根据这两种必然性，我们可以得到两种可能性：如果 S 的第一内涵在情形 W 为真，则我们说 W 确证 S，即 S 在认知上可能，这种可能性被称为第一可能性；如果 S 的第二内涵在世界 W 为真，或者说世界 W 被视为反事实的，从而令 S 为真，则我们说 W 满足 S，即 S 形而上可能，查尔莫斯把这种可能性称为第二可能性。

如果查尔莫斯的定义方式是可行的，那就在认知领域和模态领域之间建立起了联系，我们就可以说 S 是可想象的，它在认知上是可能的，反之亦然。因为 S 并没有被先天排除，虽然还不能直接得出可想象的就是可能的，但查尔莫斯认为在可想象性和可能性之间仍然存在某种关联，通过对可想象性概念和可能性的分析，我们能够从有关可想象性的前提得出形而上可能世界的结论。

三、二维语义方法中的 CP 论题

首先，查尔莫斯对 CP 论题的反例进行了深入分析。他把反例分为两类：概念上混淆的和模态上混淆的。对于前者，很显然，并非所有可想象的都是可能的并不意味着所有可想象的都是不可能的，也就是说，可想象性和可能性并非完全没有联系，这需要对可想象性这一概念进行分析。

一方面，查尔莫斯区分了当下的可想象性（*prima facie conceivability*）和概念的可想象性（*ideal conceivability*），前者受到当下主体认知能力的限制，而后者依赖于概念上的理性反思。比如，“ $2+2=5$ ”显然在这两种意义上都是不可想象的，复杂的数学真理的错误概念上不可想象，而当下可想象（如果主体对相关知识一无所知），而一只会飞的猪无论如何都是难以想象的。另一方面，查尔莫斯又区分了否定的可想象性（*negative conceivability*）和肯定的可想象性（*positive conceivability*）。当 S 不能通过先天推理来排除时，我们就说 S 在否定意义上是可想象的。如果主体能够毫无矛盾地设想 S 如此这般的条件，我们就说 S 在肯定意义上是可想象的。

在查尔莫斯看来，相比当下的可想象性，概念的可想象性更适合用来与可能性建立联系，因为前者显得过于宽泛。比如，当我们说哥德巴赫猜想和这一猜想的反面都是可想象的，这其中的可想象性就是指当下否定的可想象性，两种情形不可能同时都是概念上可想象的，也没有理由认为概念的可想象性蕴含可能性存在反例。

其次，查尔莫斯用二维语义方法澄清模态上的混淆。以“水不是 H_2O ”为例，根据二维语义，当我们想象水不是 H_2O 时，我们想象了这样一个条件，即 XYZ 而不是 H_2O 充满海洋湖泊的孪生地球，没有理由怀疑这样一个 XYZ 世界（孪生地球）是形而上可能的，而如果 XYZ 世界成为现实的，我们就必须承认水是 XYZ。因此，尽管 XYZ 世界并未满足“水不是 H_2O ”（因为 XYZ 世界不是现实的），但 XYZ 世界确证了“水不是 H_2O ”（因为 XYZ 世界是形而上可能的）。也就是说，“水不是 H_2O ”虽然不具有第二可能性，或者说形而上可能性（因

为其第二内涵在所有世界为假),但在某些情形中是可能的,因而具有第一可能性。

就“可想象性”的某种意义来说,查尔莫斯认为,水不是 H_2O 是不可想象的——如果水在现实世界毫无疑问地就是 H_2O ,在这种意义上,任何水不是 H_2O 的可想象性条件(如孪生地球)更应被描述为水仍是 H_2O 的可想象性条件,只是其中的水状物质不是 H_2O 。因此“水不是 H_2O ”看似可想象,其实并非真的可想象,查尔莫斯把这种可想象性称为第二可想象性。第二可想象性往往是后天的,因为“水是 H_2O ”这一点我们只是通过后天经验才知道的。与此相对应的是涉及先天领域的第一可想象性。 S 是否具有第一可想象性取决于先天推理,否定的可想象性也是一种第一可想象性,因为它的定义方式是看什么东西能够被先天地排除,而“水是 H_2O ”并不是先天建立起来的,在这种意义上我们可以正确地说“水不是 H_2O ”是可想象的。

因此,当我们说“水不是 H_2O ”这类句子可想象但不可能的时候,我们是在第一可想象性的意义上使用“可想象”,而在第二可能性的意义上使用“可能”,第一可想象性当然不蕴含第二可能性。但我们可以说“水不是 H_2O ”具有第一可想象性和第一可能性,这是否意味着第一可想象性蕴含第一可能性?查尔莫斯并没有直接给出这种蕴含关系成立的依据,而只是指出我们找不到这种蕴含关系的反例^①,因而第一可想象性蕴含第一可能性是无法拒斥的。尽管这样的论证并不充分,但已经足够在认知领域和模态领域之间建立起联系。

结合上面对可想象性概念的分析,对传统CP论题进行修正后就产生了两个精确的陈述^{[5][17]}:

(CP+) 概念上肯定的第一可想象性蕴含第一可能性。

(CP-) 概念上否定的第一可想象性蕴含第一可能性。

由于否定的可想象性蕴含肯定的可想象性,因此CP-蕴含CP+,但反过来是否成立似乎就很难确定了,因此CP-或许比CP+更强。但这种区别对可想象论证并没有太大影响,因为在查尔莫斯看来,怪人具有概念上的第一可想象性(无论是否定的还是肯定的)是显而易见的,按照修正的CP论题,怪人也具备第一可能性。不过这时候怪人仍然不具有形而上的第二可能性,但解决这一点并不困难,因为,根据二维语义,如果表达式 S 的第一内涵和第二内涵相同,则一个世界确证 S 当且仅当它满足 S ,也就是说,只要 S 具备第一可能性,它同时也就具备了第二可能性。为了解决这些关键问题,查尔莫斯对传统的可想象性论证进行了修正。

四、修正的可想象性论证

通过对可想象性和可能性的二维语义分析,查尔莫斯解决了CP论题的合理性,进而对传统的可想象论证进行了改进,论证过程如下^{[6][115-119]}:

- (1) 怪人具有概念上肯定的第一可想象性;
- (2) 概念上肯定的第一可想象性蕴含第一可能性;
- (3) 怪人具有第一可能性(由(1)和(2)得出);
- (4) 物理术语的第一内涵和第二内涵要么一致要么不一致;
- (5) 现象术语^②的第一内涵和第二内涵是一致的;
- (6) 如果物理术语的第一内涵和第二内涵一致,则怪人具有第二可能性(由

^① 查默斯承认可能存在潜在的反例,但是那会导致一种“强必然性”。根据二维语义的分析,像“水是 H_2O ”这种后天必然性的例子都具有必然的第二内涵和偶然的第一内涵,也就是说,这些陈述具有形而上的必然性而不具有认知上的必然性,这种必然性可以成为“弱的后天必然性”;而如果这种后天必然命题的第一内涵也是必然的,那么这种必然性就是一种“强的后天必然性”,可以简称为“强必然性”。如果否认第一可想象性蕴含第一可能性,势必会导致这样一种强必然性,而这种强必然性很难得到辩护。

^② 即关于意识现象的术语,这里“现象的”相对于“物理的”而言。

(3) 和 (5) 得出)

(7) 如果物理术语的第一内涵和第二内涵不一致, 则怪人不具有第二可能性, 因而现象事实被更深层的实在所蕴含。

(8) 要么怪人形而上可能, 要么现象事实被更深层的实在蕴含 (由 (4)、(6) 和 (7) 得出)

(9) 如果怪人形而上可能, 则物理主义错误。

(10) 如果现象事实被更深层的实在蕴含, 则物理主义错误。

结论是物理主义错误 (由 (8)、(9)、(10) 得出)

这个论证的前提 (2) 即修正的CP论题, 前提 (4) 是一个分析命题, 至于前提 (5), 涉及克里普克对有关感知特性的概念的理解。克里普克认为像“疼痛”等感知特性是我们意识中最简单最直接的, 我们无法区分疼痛本身和对疼痛的感觉, 而对于物理概念我们能够区分其属性与实在。查尔莫斯采纳了这一观点, 认为现象术语的第一内涵和第二内涵是一致的, 这一点并不会引起什么疑义^①。因而关键在于阐明: 当物理术语的第一内涵和第二内涵一致或不一致时, 将分别导致何种结论。

假定它们相同, 根据二维语义, 如果句子 S 的第一内涵和第二内涵相同, 则一个世界只有在满足 S 的情况下才能确证 S, 也就是说 S 只有具备了第二可能性才能具备第一可能性, 既然 P 和 Q 的第一和第二内涵都一致, 那么“P&~Q”具备了第一可能性就必定具有第二可能性, 因而怪人是形而上可能的, 物理主义是错误的。

假定它们不同, 这时候我们无法推出怪人是形而上可能的, 而只能满足于怪人的第一可能性, 但同样也会产生物理主义者无法接受的结果。物理主义坚持世界在微观层次上由基本粒子构成, 如果物理术语尤其是微观物理学术语的第一和第二内涵并不一致, 那就意味着微观物理概念的指称通过特定的“理论角色”(即结构特征上的规定性) 得以确定, 但却指向这些角色背后的内在本质, 也就是说, 微观物理概念的第一内涵指向任意充当该角色的属性 (不涉及内在本质), 而第二内涵指向实际担当该角色的属性 (也就是得到例示的属性, 而不涉及其角色本身)^{[7]179}。但是, 既然怪人只具有第一可能性, 那就意味着“P&~Q”的第一内涵在某情形为真, 而第二内涵在所有世界为假。也就是说, 必定有某世界 W (如怪人世界) 确证 P 且确证 ~Q, 而没有世界满足 P 且满足 ~Q。这样的话, 世界 W 与我们的世界显然就是同构的 (理论角色或者说微观物理层次上的结构相同), 区别在于世界 W 中的“理论角色”由不同的内在属性来担任。这样问题就来了, 我们的世界蕴含 Q, 而物理上与我们的世界同构的世界 W 却不蕴含 Q, 那么 Q 就不是由结构特征决定的, 只能由不同的东西 (即内在的属性) 来蕴含, 也就是说, 意识或感知特性萌生于内在的本质属性, 这些属性的本质并不是通过感知或科学揭示给我们, 而又与意识紧密地联系在一起, 它们本身可能就是感知特性或元感知特性 (即通过一定方式结合在一起从而构成感知特性)。查尔莫斯称这种观点为元泛心论 (panprotopsychism)^②, 如果接受这一观点, 那么, 当我们想象怪人的时候, 我们在现实世界的物理系统中所确定的只

^① 查尔莫斯指出, 前提 (5) 对于这一论证并不是必须的, 即使现象属性的第一内涵和第二内涵并不一致, 也能得出物理主义错误的结论, 论证过程大致如下: 首先假定物理术语的第一内涵和第二内涵一致, 根据物理主义, “P→Q”是后天必然的, 按照二维语义的分析, 这意味着“P→Q”的第一内涵是偶然的; 令 S' 的第二内涵和 S 的第一内涵相同, 则“P'→Q'”的第二内涵, 即“P→Q'”的第二内涵是偶然的, 因而“P→Q'”表达命题也是偶然的, Q' 当然为真, 但是根据物理主义, 对于所有真句子 S, “P→S”是必然的, 因而物理主义错误。而如果物理术语的第一内涵和第二内涵不一致同样会导致元泛心论, 因此物理主义错误。

^② 由于这一观点与罗素的“中立一元论”十分接近, 查尔莫斯也径直称之为“罗素一元论”。

是结构属性，而非内在的元感知特性。显然，这种具有二元论倾向的观点很难被物理主义者接受，而如果把这些内在属性也视作物理属性又显得十分勉强^①，毕竟物理学对于这些内在属性究竟是什么没有任何有效的说明。

可见，不论物理术语的第一和第二内涵是否一致，其产生的后果都是物理主义者无法接受的，因为如果一致，那么怪人就是形而上可能的，如果不一致，那就会导致元泛心论。就论证过程来看，这种基于二维语义的可想象性论证非常严密，关键在于查尔莫斯通过二维方法使得 CP 论题更加精致，更加容易辩护。但这并不意味着这种修正的可想象性论证解决了所有难题。

五、二维语义论证的辩护

查尔莫斯基于二维语义的可想象性论证引起了广泛关注，其中所涉及的一些关键问题——怪人是否可想象、CP 论题是否成立、二维语义本身是否有效等，都颇受争议。查尔莫斯针对这些反对意见为自己的观点进行了辩护。

关于怪人是否可想象问题，巴罗格 (Balog, Katalin) 提出了一个有趣的反对意见，其策略大致如下：由于怪人世界是我们世界的精确的物理摹本，那么物理主义在怪人世界显然是绝对正确的，而怪人哲学家同样能够在他们的世界提出一个查尔莫斯式的可想象性论证以证明物理主义者在他们的世界是错误的，如果查尔莫斯的论证过程明确无误，那么这一矛盾就说明怪人论证的前提出了问题^{[8]510-514}。而查尔莫斯辩护说，原来的论证有效并不代表怪人哲学家的论证也有效，原先的论证前提中的 Q 指的是某人具有意识现象，但怪人哲学家的论证中的 Q 相应的也是指“某人具有意识现象”吗？而如果怪人世界的两位哲学家就意识现象或感知特性是否存在这一问题发生争论，哪一种观点应该得到支持呢？正如普特南所指出的：“可能世界中的人们虽然能够想和‘说’和我们所能想、所能说的任何话语，但……他们不能指称我们所指称的东西。”^[9]原来论证中正确的前提在怪人哲学家的论证中是错误的。

我们认为查尔莫斯的辩护是有道理的。巴罗格的反对意见颇有诡辩的意味，虽然给我们很多重要的启示。怪人世界作为一种认知“情形”在某种意义上（先天无法排除）是可想象的，或者说它具有概念的否定的可想象性，但是否也具有概念的肯定的可想象性呢？毕竟后者要求我们必须概念上毫无矛盾地构想使怪人世界如此这般的条件，也就是说，我们还必须为条件填充细节。这其中就有两个问题：首先，我们需要在多大程度上填充这样的细节？丹尼特认为，那些接受怪人可能性的人并没有充分地设想它们：“当哲学家们宣称怪人可想象时，他们总是低估了概念（或想象）的艰巨性，所想象的东西最终与他们原先的规定相违背。”^{[10]322} 而当我们充分地设想怪人的时候，立刻就会出现第二个问题——这种概念上的无矛盾的构想是否果真是完全无矛盾的？先天性并不等于概念上的一致性，比如，“猫是动物”在概念上是协调一致的，但却并非先天的，“一个特定的类别是不是动物的一个种这样的问题，是一个要求经验研究的问题”^{[11]101}。概念分析不可能完全摆脱经验知识的背景，当我们运用部分概念构想怪人以及怪人世界的时候，这些概念之间或许是协调的，但在更广泛的概念背景中，这些概念能否获得其一致性就很值得怀疑。我们知道，在能量守恒和转化定律创立之前，人们认为机械永动机在概念上是协调的，而在热力学第二定律创立之前，人们又认为热能永动机在概念上是协调的。因此，概念上的一致性并不能保证怪人在肯定的意义上是可想象的。当然，诚如克里普克所指出的，当我们谈论可能世界时，我们“只能部分地描述可能世界”，完整地描述

^① 斯图加就要求扩展物理主义的内涵，他认为怪人是否可想象取决于我们如何理解物理属性，如果将这些内在属性也视为物理的，则怪人既不可想象也不可能。见：Stojar, D., 2001, 'The conceivability Argument and Two Conceptions of The Physical', *Philosophical Perspectives*, 15:393-413.

可能世界的全部细节既不必要也不可能，这一点即使在现实世界也不可能做到。但是不能完整地描述可能世界并不意味着我们不需要充分设想它们，提出一种假设比让我们相信它可能为真容易得多。巴罗格的反对意见实际上体现了我们对怪人的无知，事实上，我们无法想象当一个物理层次上与我们一模一样但没有意识经验的孪生人站在我们面前时将是何等的情形，我们也无法预知这位孪生人的任何行为、动作、表情等等。可以说，怪人的可想象性源于我们的直觉，但最终与我们的直觉相悖。

关于 CP 论题是否成立问题，我们发现查尔莫斯对第一可想象性和第二可想象性的区分其实很取巧，它们分别与第一可能性与第二可能性密切相关，查尔莫斯强调关于可想象性的讨论只能在陈述的层面进行。因此，一个陈述的可想象性实际上涉及两个方面：一是相关可能世界的可想象性；二是陈述在该世界中的真值。当查尔莫斯说陈述 S 具有第一可想象性，实际上是说，S 在想象的情形 W 中为真，而情形 W 本身是形而上可能的，或者说可想象世界与可能世界至少是共外延的，于是从可想象性推出可能性就是顺理成章的事情了。

这里有两点值得注意，首先，与可想象性概念的前两种区分不同，第一可想象性和第二可想象性的区分一开始就和两种可能性的区分联系在一起，它们的定义方式甚至是一样的，我们说 S 具有第一可想象性，如果它在某情形中得到确证，也就是说 S 的第一内涵在该情形中为真，按照第一可能性的定义 S 当然也具有第一可能性。显然，第一可想象性蕴含第一可能性是预先规定好的，而不是通过证明给出的。事实上，查尔莫斯对这一关键问题也从来没有给出正面的论证，而只是反复强调反对这一蕴含将导致一种无法辩护的强必然性。其次，查尔莫斯在这里有一个巧妙的转换，陈述 S 的可想象性变成了可能世界本身的可想象性，而后者是可能世界语义学的根本，因而，查尔莫斯说，即使克里普克也不否认“水不是 H₂O”在认知上是可想象的。但普特南认为，我们对《命名与必然性》存在一种误读，克里普克实际上并不赞成说“水不是 H₂O”是可想象的，而只是说可以想象“水不是 H₂O”是可想象的。如果用“◇”表示模态算子“可想象的”，则“◇◇p → ◇p”是不能接受的^{[12]55}，虽然它没有违反模态逻辑通常的规则。这里普特南强调的是，即使我们能够想象水不是 H₂O 的孪生地球，但这不等于说水不是 H₂O 直接就是可想象的，更不能说水不是 H₂O 是可能的。

这里更进一步的分歧实际上涉及克里普克和查尔莫斯对可能世界的不同理解。克里普克和普特南坚持认为可能世界是“由我们赋予的描述条件给出的”^{[13]23}，并不具有本体论的实在性。按照这一观点，像查尔莫斯所说“当某情形或可能世界成为现实的（或被视为现实）”是不恰当的，是对可能世界的误解，这样，克里普克的语义学考察本质上只能是一维的。很难说对于可能世界的理解哪一种更正确，但是取消现实世界的特殊地位更便于一般地讨论表达式在任意可能世界中的意义，因此查尔莫斯的二维方法对于解决语义学中长期以来的一些争论会很有帮助，并且在二维框架中我们确实很难拒绝修正的 CP 论题，尽管查尔莫斯并没有给出 CP 论题的直接肯定性的证明。

查尔莫斯可想象性论证中的另一个重要问题是它预设了感受性的存在及其与物理属性的根本区别，而且感受性不能概念地还原为物理属性或功能属性，因而这个论证攻击的是二者之间的必然蕴含关系。然而，我们关于感知特性的信念除了诉诸直觉和内省并不能得到有效的说明，因此感知特性及其相关概念本身就引起了极大的争议。查尔莫斯这一坚定立场最终通向一种他所谓的“自然主义的二元论”，也就是元泛心论，它与后天物理主义的观点有着实质的区别，但似乎也还没有完全跳出还原主义的窠臼，有学者将查尔莫斯的这种观点归结

为一种“功能的还原主义”^{[14]99-100}，这显然并不恰当，因为感受性不是由微观物理世界的结构属性决定，而取决于填充这些结构的内在属性，这是一种本质还原，甚至是还原为一种更深层次的实在。显然，这种较强二元论倾向的观点并不容易被接受。

六、结束语

通过可想象性论证彻底驳倒物理主义并不容易，意识问题能否摆脱形而上学而获得完全科学上的解释仍处在激烈的争论之中，但有一点很清楚，即使科学的解释取得成功，我们也不必担心现象术语会被取代，正如我们即使知道了水的分子式仍然称那种液体为水而不是 H₂O。虽然查尔莫斯的可想象性论证存在缺陷，但他的认知二维语义分析却颇为新颖，引起了极大关注。在二维语义的框架内，查尔莫斯恢复了先天性与必然性、认知领域和模态领域之间的联系，强调意义与认知的普遍关系，为传统的描述理论进行了辩护，同时又不排斥因果指称理论，这样一种尝试确实非常吸引人^[15]，而通过语义分析为反物理主义论证提供支持，也体现了心灵哲学和语言哲学之间越来越来紧密的联系。

参考文献

- [1]休谟.人性论.关文运 译, 商务印书馆, 1980.
- [2] Brueckner, A.Chalmes's conceivability argument for dualism. *Analysis* 61,2001:187-193; Worley, S. Conceivability, Possibility and Physicalism. *Analysis*, 63, 2003.
- [3] Putnam, H.Is Water Necessarily H₂O? In his *Realism With A HumanFace*. Cambridge, Massachusetts: Harvard University Press, 1990.
- [4]Chalmers,D.J.The Foundations of Two-Dimensional Semantics. In M. Garcia-Carpintero & J. Macia (eds.), *Two-Dimensional Semantics: Foundations and Applications*, Oxford University Press, 2006.
- [5] Chalmers,D.J.Does Conceivability Entail Possibility? In T. Gendler & J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford University Press, 2002.
- [6] Chalmers,D.J.Consciousness and its Place in Nature. In S. Stich & F. Warfield (eds.).*Blackwell Guide to Philosophy of Mind*. Blackwell, 2003.
- [7] Chalmers,D.J.Does Conceivability Entail Possibility? In T. Gendler & J. Hawthorne (eds.).*Conceivability and Possibility*. Oxford University Press, 2002.
- [8] Balog, K.Conceivability, Possibility, and the Mind-Body Problem *Philosophical Review* 108, 1999:.
- [9]普特南.理性、真理和历史.童世骏、李光程译.上海译文出版社, 2005.
- [10] Dennett, D. C. The Unimagined Preposterousness of Zombies. *Journal of Consciousness Studies* 2, 1995.
- [11]克里普克.命名与必然性.梅文译, 上海译文出版社, 1998.
- [12] Putnam, H. Is Water Necessarily H₂O?In his *Realism With A HumanFace*.Cambridge, Massachusetts: Harvard University Press. 1990.
- [13]克里普克.命名与必然性.梅文译, 上海译文出版社, 1998,
- [14]曹向阳.当代意识科学导论.东南大学出版社, 2003.
- [15]黄益民,《二维语义学及其认知内涵》,载《哲学动态》2007(3)。

On Chalmers' Conceivability Argument Based on Two-dimensional Semantics

Wei yi-dong Chen Jing-kun

Abstract: In the contemporary philosophy of mind, a number of popular arguments against physicalism conclude that there is an ontological gap between physical process and consciousness as the epistemic gap between physical truths and

phenomenal truths, but physicalists hold that the epistemic premises do not entail ontological conclusion. By virtue of two-dimensional semantics and rebuilding the bridge between conceivability and possibility, Chalmers improves the traditional conceivability argument and argues that the inference from epistemic premises to ontological conclusion is reasonable. This is a real challenge for physicalism, but this paper argues that, as some deficiencies, the improved conceivability argument cannot confute physicalism virtually.

Key words: Chalmers, conceivability Argument , two-dimensional semantics, physicalism