



首页 → 学术文章 → 应用伦理学基础理论

## 赵汀阳：论道德金规则的最佳可能方案

**提要：**道德金规则是伦理学中公正原理的一个浓缩表达。本文分析了传统的道德金规则的典型民间版本和学术版本，指出传统金规则要得以成立，在逻辑上至少需要两个条件：人际共识和价值共识。而在今天，价值共识已经不存在，所以传统金规则不再成立。作者试图通过从“主体观点”到“他者观点”的转换而提出一个关于道德金规则的新版本，即以“人所不欲，勿施于人”代替“己所不欲，勿施于人”，并试图论证其逻辑上的完美性。本文中的核心论点最早于2000年提出，由于该问题涉及到生活基本准则，事关重大，因此经过听取多方意见和反思，在这里进行了更严格的论证和修改。 (<http://www.tandem.cn>)

### 1. 问题背景

首先需要提及为什么要讨论道德金规则的问题背景。

一般地说，金规则指的是能够概括地表达一个伦理体系的总精神的一条道德原则，也就是能够“一以贯之”的普遍原则；如果在学术意义上说，它就是伦理体系的一个元定理，它是对伦理体系中各种具体规则的总指导和解释。金规则总是非常稳定的，除非社会发生巨大变化，否则它不会变化，而现在正是一个巨变的时代。 (<http://www.tandem.cn>)

全球化把以前不明显的许多问题变成了明显的问题，其中一个典型问题就是各种文化/知识体系之间的关系。亨廷顿关于文明冲突的论点虽然有着根本性的错误，但它却揭开了“对话/交往”问题的真正底牌。自苏格拉底以来，理性对话就被认为是通向普遍承认的真理之路，到今天，哈贝马斯还坚持认为，完全合乎理性标准的正确对话必定能够产生一致认可的理解。但是哈贝马斯忽略了一个关键性的问题底牌，这就是，理解不能保证接受[i]。理性对话有可能达到一致的理解，但是人们想要的不仅仅是被理解，而是被接受。接受才是“对话/交往”问题的终点，这一危险的底牌在以前的哲学分析中被有意无意地回避了。达成共识和合作的充分理由不是互相理解，而是互相接受。可是互相接受的问题超出了知识论和理性所能够处理的范围。显然，“接受问题”迫使知识论上的“主体间”问题深化为实践或价值理论上的“人际”问题，如果扩大计算单位，则成为“文化间”问题。人们在考虑知识时必定同时考虑价值，知识问题和价值问题是共轭的。这样就回到了哲学的正宗模式上了，无论希腊还是中国先秦，知识问题都是从属于伦理/政治问题的。 (<http://www.tandem.cn>)

当把“接受问题”计算在内，“对话”就变成了“对待”，伦理学就成为第一哲学的一个部分（列维纳斯甚至相信第一哲学只能是伦理学[ii]），另一个部分非常可能是政治哲学（施米特相信政治生活是最基本的生活形式，而政治问题就是区分敌友[iii]）。伦理学和政治哲学的一个最基本问题就是“如何对待他人”。列维纳斯非常正确地论证了他人的绝对性，他人是一个无论如何无法被“我”的主观性所消化的外在绝对存在，主观性（subjectivity）化不掉他者性（otherness），所以他人超越了我的主观性，是我的生存条件和外在环境。他人会反抗，他可以不合作，所以超越了我，所以他人是我需要对待的最严肃的问题。今天人们特别感兴趣的全球合作、全球共识、全球价值之类，在学理上都依赖着关于“他人”的理论。对待他人的总原则在伦理体系中就表现为所谓的“金规则”。 (<http://www.tandem.cn>)

各种文化的伦理体系中都有着至少一个被认为在理论上无懈可击、在实践上历久常新因此非常可能是万世不移的道德原则，它被当作是一个伦理体系的基石。通常人们把这类据信为“绝对无疑的”道德原则按照基督教伦理的习惯称为“金规则”（the golden rule）。根据孔汉思和库舍尔的研究[iv]，据说不仅各种文化中都有金规则，而且这些在历史中各自独立地自发生成并且以不同方式表述出来的金规则在含义上“都惊人地相似”，几乎可以说其逻辑语义是完全一致的。于是他们认为，这种一致性表明了金规则是放之四海而皆准的普遍必然原则。 (<http://www.tandem.cn>)

可是由经验巧合去推论普遍必然性，这在理论上说（根据休谟定理），是不可以接受的，因为不存在这样一种逻辑，或者说不存在这样一个必然有效的推论模式。在理论上，道德金规则一直没有被成功地证明。而在实践上，虽然金规则一直在伦理体系中有着最重要的意义，但也一直存在着一些根本性的困难，这些困难在古代社会里也许不很明显，可是在当代社会里就变成了严重挑战，比如说，在多文化共享的社会空间里，各种文化（各种文明、宗教、传统

和政治理想)甚至各种亚文化(女性主义、环保主义、同性共同体等)都拥有不同的价值观,这意味着它们在“想要的和不要的”东西上有着不可通约的需要和评价标准,“己所不欲勿施于人”这样的金规则在这里已经没有能力处理那些价值问题了,至少可以说,在许多事情上都会遇到严重挑战。在这个意义上,基于目前被普遍承认的金规则的所谓“全球伦理”是不可能成功的,因为这一全球伦理运动仅仅考虑到金规则在“空间”中的普遍存在,而没有考虑到社会关系在“时间”中的变化。从空间的角度去看,人们似乎都承认有那么一种古老而普遍的金规则,但是,问题已经在时间中发生了巨变,问题变了,答案自然也应该有所变化,传统金规则格式不再是无懈可击的了。

## 2. 人际共识和价值共识

如上所述,金规则表达的是如何对待他人的人际共识。传统金规则有许多版本(在后面我们再作分析),但无论什么版本,其根本精神是完全一致的,或者说,它们的逻辑语义是同样的。传统金规则的基本假定是所有人(或至少大多数人)具有价值共识,也就是所谓“人同此心,心同此理”。与这个基本假定相配合,其方法论则是“推己及人”。以此假定和方法论就必然得出传统金规则。

问题正在于此:诸如“己所不欲勿施于人”这样的人际共识是有效的,当且仅当(iff),一个社会具有共同价值观,也就是具有关于“想要的/不要的”的价值共识。这两个共识必须同时存在,否则传统金规则就不可能成立。为什么古典的金规则在今天遇到困难了?其秘密就在于现代社会失去了价值共识,因此原来的人际共识就失去了得以成立的必要条件。

这个问题在古代社会所以没有出现,是因为任何一个古代社会都还没有发展出许多互相冲突又几乎同样有力的价值观,即使人们在价值问题上有某些不同意见,也还没有形成各种同样有影响的权力话语,就是说,在古代社会,人们想要的和不要的基本一致或者说大同小异,即使有些另类人物的奇谈怪论,也只是一些学术性的观点而没有成为有社会影响力的话语,没有成为主流,因此不影响社会的总体价值选择。比较粗略地说,古代社会的冲突的主要原因是利益问题而不是价值观问题。尽管古代也有价值观冲突,但基本上只是学术现象。价值观冲突成为社会现象是现代社会的产物,是所谓启蒙的产物,是自由和平等的结果。很容易想象,在价值观基本一致的社会条件下,就有了“人同此心,心同此理”的普遍现象,于是“推己及人”的方法论就能够适用,金规则的古典版本就是顺理成章的了。可是今天不再如此。

当然,必须承认,即使在价值观基本一致的古代社会里,例外情况总是有的。但是,更应该强调的是,例外的社会现象对于社会一般知识不能构成挑战。这是个有趣的知识论问题,可以称作“例外”的知识论问题。对于自然科学尤其是逻辑知识来说,如果万一出现了“例外”,就是无比严重的问题,因为“例外”构成了挑战普遍必然性的“反例”。可是对于社会生活来说,“例外”是软弱无力的,因为“例外”无法构成对主流价值或者统治性话语的挑战,往往可以忽略不计,“反例”对于社会知识来说是非常可笑的。在古代社会里虽然总有某些例外的价值观点,但被主流价值观所淹没。所以可以说,古代社会具有价值共识。

作为传统金规则的必要条件的价值共识是在现代被破坏的。事实上,从现代开始以来,传统金规则就注定要出现问题了。从价值观念方面去看,现代开始于“平民”反对“贵族”的价值观以及相应的制度安排,自由和平等的要求注定了价值观的多元化和冲突。现代社会的产生当然有着各种各样的重要原因,不过其中的价值观革命可能是最深刻的。尼采可能最早意识到现代性意味着一种彻底革命的价值观,他指出现代就是奴隶反对主人的运动,当然也就要用属于奴隶的“低贱的”价值观去反对主人的“高贵的”价值观。后来,列奥·斯特劳斯又指出现代性的另一个相关的基本精神是“青年反对老年”,也就是今天反对古代——青年被用作隐喻指示现代,因此,他认为“古今之争”是最大的价值观冲突。“青年”这一隐喻意味深长,它不仅可以说现代以“进步/落后”的技术指标替代了传统的“好/坏”人性标准,以“新/旧”的时尚指标替代了传统的“卓越/拙劣”的品质标准,而且还因此导致了无法止步永不停息的“推陈出新”运动,从积极的方面看,这是过不完的青春期,从消极的方面看,这又使得精神积累不再可能。这就是现代性。

在这里我们关心的不是关于现代性的批判,而是想说,现代性这种“新/旧”和“进步/落后”的价值指标必然形成各种各样价值观的大量生产和互相冲突,因为各种价值观都有理由以“新”和“进步”作为其合法性根据。这样来看,现代性就是没有一致价值观的时代。人的解放导致思想解放,思想解放导致价值多元。正如前面所分析的,一旦失去“价值共识”这个基础,传统金规则就会失去普遍有效性。其实不仅仅是金规则,几乎所有古典的标准都因为现代社会的各种新价值(平等、个人主义、进步、新奇、数量化、多元化等等)而失去效力。金规则问题可以看作是现代问题的一个典型案例,它表明在这个彻底现代化的时代,各种原来认为的普遍原则也都不得不更新换代。

## 3. 对等性结构和互换性结构

道德金规则，无论什么样的版本，都意味着伦理体系的一个元定理。我们知道，任何一个规则系统（哥德尔意义上的严格系统），如果其内容足够丰富的话，就必定是不完备的。而伦理系统本来就是而且只能是很不严格的系统，而且其系统内容必定极其丰富，因此漏洞百出就不足为奇。按照我在《论可能生活》里的论证，一条伦理规则要应用的情景几乎是无穷多的，而生活情景不可能完全一样，因此任何一条规则总是不得不根据具体情景被灵活解释。这种规则与实践的差距就难免导致“标准失控”的难题。于是，要维持伦理系统的解释在大体上的稳定性（绝对严格的稳定性是不可能的），就必须有一些明显普遍有效的一般理念来对各种具体规范作出最后的解释和判断。也就是说，生活的可能情景无穷多，而且其变化情况无法预料，伦理规范本身又总是含糊的（例如“不许说谎”就是一个含糊表达，而如果说成“无论何时何地都对任何人都决不说谎”就会遇到严重困难），因此需要有某些能够应付“所有情况”的基本理念，也就是能够“以不变应万变”的原则来给出最后解释。这些基本理念就是幸福、公正、自由的理念（有的伦理体系还要求更多的基本理念）。对这些根本性的理念的表述可以是一些复杂的理论原则，同时往往也表述为清楚明白的实践原则，这些原则就成为伦理体系的元定理，也就是有能力对某个规则体系的总体性质进行反思并且作出判断的“最后”定理。

通常所谓的金规则正是表达着公正的理念，因为无论在理论上还是实践上，公正原则都是对“如何对待他人”这一问题的唯一理性回答。公正原则对任何涉及他人的行为规范做出理性的判断和解释，因此意味着能够普遍承认的人际关系原则。

在传统金规则的各种版本中，最有名的也可能是最典型的是基督教的金规则和孔子原则。基督教金规则的正面表述是：“你若愿意别人对你这样做，你就应当对别人也这样做”；其反面表述则是：“你若不愿意别人对你这样做，你就不应当对别人这样做”。孔子的正面说法是：“己欲立而立人，己欲达而达人”，其反面说法是：“己所不欲，勿施于人”。

学院派哲学家对这些民间风格的表述不太满意，于是又有了一些学术版的表述，最著名的是康德版，即他的道德普遍律令：“你只能按照你希望能够成为普遍规律的行为准则去行为”[v]。康德认为只有他给出的这个伦理原则才是真正严格的，因为它不需要利用实践经验，仅仅通过理性本身而获得的。基于民间经验的金规则被认为应该以基于理性本身的原则为准去重新理解。还有比康德版更细致一些的西季维克版（但未必比康德版高明）。西季维克指出，金规则的正面表述肯定是错误的，因为人们完全可能愿意互相帮助做坏事。金规则的反面表述虽然不是错的，但却仍然不准确，西季维克的修改版大概是这样的：“对于任意两个不同的个体，A与B，如果他们各自情形上的不同并不足以成为在道德上加以区别对待的根据，那么，如果A对B的行为不能反过来同时使B对A的同样行为同样是正确的话，这一行为就不能被称为在道德上正确的行为”[vi]。如此等等。尽管哲学家们相信他们的学术版比民间版严格得多，但仍存在着严重问题。真正根本的问题并不在于什么样的表述更为严格，而在于所有传统版本，无论是民间版还是学术版，都有着同样错误的思想出发点。学术版尽管严格，却只不过严格地继承了民间版的错误。其中比较好的版本应该是孔子的正面表述，尽管正面表述的金规则比较冒险，容易出现明显的漏洞，但孔子的正面表述包含了“立”和“达”这样的模糊概念，就多少回避了困难，但基本的困难仍然存在。

不管那些学术版的金规则如何改进表述上的逻辑性和严格性，但各种金规则的基本结构仍然是一致的，即强调互相对待的对等性（reciprocity）。这个结构没有错误。对等性结构应该是表述“我与他人”的关系的最合理结构，因为必须有对等性结构才能够表达公正关系，而金规则就是为了表达公正关系的。在《论可能生活》（1994年版；2004年版）中，我也是以“对等性”为基本结构来分析和定义公正的（在定义公正时，一部分哲学家主要强调“对等性”，另一些哲学家则主要强调“公平性”，这是两个众望所归的基础，但在本质上说，“对等性”是更为重要的因素，只有当承认了对等性，公平性才有意义。对等性固然不错，但事情没有那么简单，我们还需要警惕另外某些可能被漏掉的因素。

公正在形式意义上直接就意味着一种对等性。人们早就意识到公正是一个“恰如其份”的概念，它意味着各得其所、各得所值。无论对于人际关系还是事际关系，公正的对等性首先表现为“等价交换原则”，即某人以某种方式对待他人，所以他人也以这种方式对他，或者某人以某种东西与他人交换与之等值的東西。这一原则虽然非常“清楚明白”，就像笛卡儿所推崇的真理那样，但它实际有效的情况却很有限，因为，只有当双方在某种情境中具有几乎同等的自由和能力时，这一原则才能够被有效地执行。于是，公正的对等性通常又表现为一些比较复杂的对等形式，比如“豫让原则”，即某人以对待什么人的方式对待我，那么我就以什么人的方式回报他（如豫让所说“……以国士遇臣，臣故国士报之”）。还有“西季维克原则”：给同样的事情以同样的待遇，而给不同的事情以不同的待遇。

然而，公正仅仅表现为对等性是不够的。怎样才真正算是对等的？这仍然是不够清楚的事情。对等原则并不能解决需要公正处理的所有问题。具体地说，等价交换原则，即A以X方式对B，因此B有理由以X方式对A，只能证明“B以X方式对A”是公正的，却无法证明“A以X方式对B”是公正的；而豫让原则，即A以(B=X)的方式对B，因此B有理由以X的方式对A，也只能证明“B以X的方式对A”的公正性，却不能证明(B=X)这一方式的公正性；同样，西季维克原则，可以表述为“按照标准X，A和B是同样的，所以给予A和B同样待遇”，也只能说明对于给定标准X，A和B得到同等待遇是公正的，却无法证明设定标准X是不是公正的，我们也就无法真正知道A和B被看成是同样的是否是公正的。

这些情况的共同问题表明，在对等性结构中，我们只能必然地证明“后发行为”的公正性，却没有理由证明“始发行为”的公正性。这是传统公正理论以及传统金规则所难以处理的问题。

为了使一个相互关系得到在场各方的共同绝对认可——排除了迫于条件、压力和强迫的相对认可——就必须使得在场各方都认可这样一个“地位互换”原则：如果A以X方式对待B是正当的，当且仅当，A认可“当A处于B的地位而B处于A的地位，并且B以X方式对待A是正当的”。对于这种地位互换关系，如果无论把“我”代入为A或B，“我”都将认可其中的行为方式，那么这一行为方式就是正当的。这个地位互换原则在利益分配上同样有效，它表现为：如果A按照X准则把A和B看成是同等的，并且A和B得到同等的利益分配是正当的，当且仅当，A认可“当A处于B的位置而B处于A的位置并且B按照X准则把A和B看成是同等的，并且因此得到同等利益分配是正当的”。这个“地位互换”原则可以看作是传统的角色互换原则的改进版。传统的角色互换主要表达的是“将心比心”或者“同情”的直观，这个直观大体不错，但是不够严格，因为无论是“我”还是“他人”都似乎是抽象的同样的人，其中暗含着“既然都是人，那么就都应该什么什么”这样的认识，这其实仍然没有超出对等性原则。事实上任何一个社会都存在着不同地位的关系，所以很难简单对等。地位互换原则就是试图发现当把地位差异考虑在内时能够产生什么样的公正关系。

现在我们可以比较完整地理解公正的意义了。首先，公正表现为对等性。这意味着允许存在着某种假设X，然后在X的基础上要求对等。这样可以保证在给定价值标准下的公正关系；其次，公正进一步表现为互换性。这意味着任何一个可能的假设X，即使它能保证对等，也必须被证明为在互换方式中是有效的。这一地位互换原则可以消除在设立价值评价标准上的不公正。

#### 4. 想象一个最佳版本

我曾经提出对金规则进行过一个本质性的修改，当然没有藐视传统智慧的意思，而仅仅是在新的问题框架中发展了传统智慧。既然金规则的反面表述被普遍认为是比较稳妥的格式，那么，我选取孔子的反面表述“己所不欲勿施于人”作为底本（包括基督教金规则在内的各种反面表述版本的语义都可以完全表达在孔子版本中），把它修改为“人所不欲勿施于人”。仅仅一字之差，但其中的人际关系发生根本性的变化。我认为这是目前所能够想象的最佳版本，它至少具有两个在理论上或者说在技术上的优势：1）假定“人所不欲勿施于人”成为替代“己所不欲勿施于人”的一种新的人际共识，它不需要以价值共识作为必要条件，而既然免除了价值共识这个苛刻要求，它因此就可以良好地适用于今天社会的价值多元情况，而且是克服由于价值多元而产生的文化冲突的一个有效原则；2）尤其在纯粹理论上，看，“人所不欲勿施于人”原则能够满足严格意义上的普遍有效性要求，相比之下，“己所不欲勿施于人”原则只是在特定条件下有效的，并非真正的普遍原则。实际上，传统金规则将成为新版金规则所蕴涵的一个特例，就是说，假如一个社会碰巧具有价值共识，在这个特殊条件下，“人所不欲勿施于人”与“己所不欲勿施于人”意义等值，但是在其他社会条件中，就只能是“人所不欲勿施于人”。就是说，“人所不欲勿施于人”的有效范围大于“己所不欲勿施于人”。

#### 5. “无人被排挤”原则

现在进一步来说明把“己所不欲勿施于人”修改成“人所不欲勿施于人”的更深入的哲学理由。

从思想语法上看，人们在思考“我与他人”的关系时一直使用的是主体观点，即以“我”（或特定统一群体“我们”）作为中心，作为“眼睛”，作为决定者，试图以“我”为准，由“我”来定义知识和标准，按照“我”的知识、话语、规则把“与我异者”组织为、理解为、归化为“与我同者”。这种自我中心的态度可能是一种自然态度，恐怕与文化无关，即使是从来都非常强调他人的重要性的中国传统文化，显然也没有完全超越自我中心的理解方式。在列维纳斯看来，这种传统的主体观点是对他人的不公。一切以我的观点为准，实际上就是对他人的否定，是实施了一种无形的暴力，一种试图化他为自己的暴力。把这种态度作为我与他人关系的道德基础显然是无效的。我们必须以把他人尊称为“您”的他人观点来代替传统的主体观点，只有以他人观点为准的理解才能尊重他人存在的“超越性”，即不会被“我”随便“化”掉的绝对性，才能避免把他人的超越性消灭在我的“万物一体化”的企图。尊重他人的超越性就是尊重他人的那种不能被规划、不能被封闭起来的无限性。

列维纳斯的观点或许有些夸张，但确实指出了问题之所在。我们无须把他人夸张为至高至尊的存在，但至少可以发现这样一个逻辑关系：尊重他人的同时就等于让自己得到尊重，因为对于别的主体来说，自己也是个必须被尊重的超越者。只有从他人观点出发才能演绎出真正公正的互相关系，即列维纳斯所谓的由“面对面”所形成的“我与你”关系。“我与你”的关系是平等的，而“我与他人”的关系则是不平等的，按照康德的说法，它是把人当作了“手段”的关系。列维纳斯这一“他人为尊”观点在思想语法上非常重要。如果从主体观点入手，心中始终以我为准，我把你贬低为在我的知识和权力范围里去解释的他人，你当然也会把我当成他人，这种似乎对等的关系逻辑地导致了实际上的互相否定，这是一种互相拆台的对等。即使好象有了互相尊重的美好意愿，但只要思维方式是主体观点，那种意愿

就无从实现（康德就希望仅仅以“善良意志”作为唯一条件，好象善良意志足够解决问题。可是良好意愿有时只不过是欺骗自己说自己是个好人，对他人并无实质好处，如果仅仅有意愿而没有行动，只能证明意愿的落空）。

根据列维纳斯的批判，我们可以有进一步的发现：所谓各种文化都认可的、据说可以作为第一普遍原则的金规则“己所不欲勿施于人”表面上似乎表达了对他人的善意，其实隐藏着非常典型的主体观点霸权，它的思维出发点仍然是“己”，它只考虑到我不想要的东西就不要强加于人，根本没有去想他人真正想要的是什么，这意味着，我才有权利判断什么东西是普遍可欲和什么事情才是应该做的，我的心灵才算是个有资格做决定的心灵，而他人的心灵和思想根本不需要在场，我可以单方面决定普遍的价值选择。这就是主体性的霸权。

取消他人的价值决定权是非常严重的错误，这里存在着深刻的权利之争，它是知识权利之争，它是关于知识的政治斗争，即谁才有权利决定什么算是合法的或正当的知识，或者说，谁才有权利决定什么算是有价值的东西以及什么算是正当的事情。知识权利不仅是政治权利，而且是最根本的政治权利。所以，即使像“己所不欲勿施于人”这样的金规则的确表达了善良意志，这种心理学的解决方式也没有能够真正解决我与他人的公正关系问题。可以说，伦理学问题的解决不能依靠心理学，无论是康德式的善良意志还是孟子式的良心，而只能指望一种政治学的解决，即必须能够使得我与他人同样拥有知识上的政治权利。

可以进一步分析，为什么主体观点在伦理学中是不合法的？我们知道，主体观点首先是一种知识论观点，是关于如何“看”世界的方法。在“看”事物时，不需要征求事物的同意，我想怎么看就怎么看。在纯粹知识论里，主体观点虽然不见得是最优的（它所欠缺的视角太多），但却是合法的。但是在其它地方，主体观点就失去了合法性。例如在交往问题中，我就不可以想怎么就怎么随便“看”别人，因为问题已经由“看”变成了“说”和“听”，其中关键还落在“听”上，我们不得不去听别人说什么，否则交往不成立，在这里，主体观点的合法性就已经很可疑了。当进入到“做”的问题，即共处或合作的问题（做任何事情都涉及他人），主体观点就完全不合法了，因为生活问题的基础是存在论而不是知识论，把知识论的原则推广到存在论中去是不合法的，显然，他人不可以被处理成像物那样的“对象”，因为他人意味着我所化不掉的外在精神，他也将按照他的精神去做事情，而且我要做任何事情都不得不与他人一起做（至少是间接地与他人一起做事情）。因此，在存在论所控制的领域里，他者性（otherness）就必定成为替换主体性（subjectivity）的最高原则。

当把他人观点计算在内，道德原则所考虑到的变量就有了巨大的变化，既然增加了必须计算在内的变量，我就有理由认为应该把金规则修改为“人所不欲勿施于人”。虽然只一字之差，但其中境界却天上地下。在“由己及人”的方法论中，可能眼界只有一个，即“我”的眼界，而“由人至人”的方法论则包含了所有的可能眼界（all possible horizons），这样才有可能尊重每个人。这个把每个人的精神考虑在内的方法也可以称为“无人被排挤”的原则。这一改变非常必要，它可以根本改变我们思考价值问题的思路 and 角度。在这里它意味着一条关于选择任何共同伦理的元规则，这个元规则蕴涵着彻底的公正，可以表述为：（1）以你同意的方式对待你，当且仅当，你以我同意的方式对待我；（2）任何一种文化都有建立自己的文化目标、生活目的和价值系统的权利，即建立自己的关于优越性（virtue）概念的权利，并且，如果文化间存在分歧，则以（1）为准。这是学术版的“人所不欲勿施于人”。

## 6. 进一步的论辩

对于我的金规则修改版（我曾经在先前的某些论文中简单地提及这个观点），有一些不同意见。目前最具学术深度的挑战是王庆节先生提出的。王庆节指出：“赵汀阳以‘他人观点’建构起来的道德金律固然可能帮助我们克服‘主体观点’，但这种以‘他人观点’为准的道德金律是否能胜任作为‘普世伦理’的‘元规则’呢？我想我们恐怕不能得出这一结论。为什么呢？让我们来看下面的例子。假设你我都是腐败的官员，而且你我都不以贿赂为耻，反而以之为荣。当我贿赂你时，我知道你想我以“贿赂”的方式对待你，并且假设你也会同意以同样的方式回报我。但是，我们知道，按这种方式进行的行为，无论是出于“主体观点”还是“他人观点”，都不能改变“贿赂”的不道德性。这也就是说，即便我对某一他人对我行为的所欲所求与此人想要得到的对待是相同的，遵循这一原则行事也不能保证永远是道德的” [vii]。

这一批评的挑战是富有启发性的，而且在某个方面来看是正确的。它试图指出任何可能设想的形式化道德原则的弱点。类似的例子曾经被用来批评康德的形式化的“普遍律令”。这是一个值得深思的问题。这种挑战与分析哲学家们喜欢的“反例”完全不同，它不是一个具体反例而是一种逻辑可能性，这样就具有力量。我们前面提到“反例知识论”问题，我相信个别反例对于社会科学知识没有挑战意义，因为社会科学所要求的知识标准不是像科学和数学那样的“无一例外”标准，它关心的是具有社会影响力的存在，个别反例无法改变任何主流的社会事实，所以个别反例对于社会科学几乎没有意义。但是，逻辑可能性则是现实的威胁，这不能不考虑。在深入分析王庆节的问题之前，可以首先考虑一个技术性问题以便削弱其挑战力度。王庆节的关于“贿赂”的假想可能性对于批评金规则的正面表述比较有效，而对于反面表述，情况则有些不同。我们所讨论的是人所“不欲”，某人A想要贿赂，这显然属于“欲”而非

“不欲”，那么，是否可以把命题“A想要贿赂”转换成命题“A不想清廉”？或许这两个命题在真值上是逻辑等值的，但却并非意义等值。显然，受贿是个经过决定而做出的行为，而“清廉”是个状态，是个还没有发生受贿事件的无为状态，一个人在没打算受贿的时候，他的本来状态就是清廉的，所以，只有“受贿”才是个欲望对象，我们似乎不能把初始状态看成欲望对象。这有点像希腊诡辩论者的例子，显然不能由我们“本来没有角”推论说我们“失去了角”。所以，“贿赂案例”只能构成对正面表述的金规则的挑战，却不太可能对反面表述的新版本金规则形成真正的威胁。

虽然以上分析削弱了“贿赂案例”的挑战性，但我们仍然必须承认，总会有办法找到某些具有挑战性的方案，不过恐怕不可能对改进后的金规则形成釜底抽薪式的威胁，但是会迫使人们承认金规则的局限性。比如说，有人喜欢战争、侵略和压迫，而和平、友好和尊重正是他所“不欲”的；又有人愿意吸毒、赌博加性变态，而改邪归正是他所“不欲”的，在这里，新版本金规则并不能提供足够理由去反对他。其实，无论什么样的金规则都有局限，就是说，任何金规则所能够治理的伦理空间是有限的，总有很大的“伦理荒地”留在外面。

康德非常可能意识到了这个弱点，这也许就是康德为什么除了形式化的“普遍律令”之外还要给出了另一条基本原则的理由。另一条原则就是甚至更为知名的“人是目的”原则。可惜这条原则的要求太高，以至于是不切实际的空话，人们不可能不在某些场合把某些人当成“手段”，否则生活不可能进行，更不可能有什么社会了。正如我在《论可能生活》里论证过的，除了单纯美好的人际关系之外，人际关系在许多必要的情况下表现为各为其利的“事际关系”，所谓“手段”的问题就必然出现，如果不把某些人当成手段，就根本做不成任何事情。不过重要的是，“人是目的”原则不再是纯形式的了，而是有价值内容的（主要是人道主义的价值观）。这样就指出了在金规则范围之外的伦理问题，也就是说，伦理学的基本问题不仅仅是公正问题，而且还有价值观的选择问题；不仅是为了解决我与他人的关系问题，而且还需要解决生活和社会的理想问题。尽管康德所选择的基本原则很有疑问，但他构造的道德原则框架确实不同凡响，难怪他可以声称他的道德原理比金规则要好得多，而且金规则必须按照他的原则进行校正。且不管人类道德体系所最需要的内容原则是什么，至少我们可以肯定，仅仅表现为形式原则的金规则是不充分的，显然还需要内容方面的约束条件。就是说，金规则不可能只是一条，而是至少需要两条，而且这两条原则是互相约束着的。

在某种意义上，康德方案既不现实又有自相矛盾之处。假定“人是目的”原则普遍成立，那么就会有一个完美到不可能出现和存在的社会。假如真有完美社会，现在我们观察到的绝大多数社会规律和规则，以及绝大多数的人文社会科学知识就失去意义，尤其是经济学、社会学和政治学的全部前提就不存在了。如果“人的目的”，人人自然就能够公正地互相对待，那条用来表达公正的形式原则就是多余的了。而既然需要公正原则，就是因为人们只能指望策略性公正，而不能指望无论何时何地都“人人成目的”的绝对公正。因此，我们有理由认为，康德的框架是伟大的，但其内在构造并不妥当，两条原则并不能很好匹配。要选择两条良好匹配的、同时又足够普遍以至于能够覆盖各种伦理学问题的“道德基本原则”确实不容易。在《论可能生活》中，我采取的也是“两条原则”的模式，但与康德所选择的原则有所不同，解释也很不同，尤其是与康德的道德原则排序不同，康德首推形式原则，然后才是价值原则，我的排序正好相反。

目前我的方案是（希望以后能够改进）：（1）幸福原则：如果一个人的某个行动是自成目的的（autotelic），那么这是一个必然产生个人幸福的行动，而如果连同这个行动的目标也是一个自成目的的事情，那么就同时成为一个能够促进人类幸福的行动。幸福既是人们的最大欲望，又是唯一对己对人都有益的行动。“幸福的帕雷托改进”是唯一具有全方位正面效果的帕雷托改进，就是说，某人的幸福不仅对自己有益，而且在任何意义上都对他人无损。相比之下，福利的帕雷托改进（即经济学意义上的帕雷托改进）虽然无损于任何他人的物质利益，但却非常可能损害他人的社会地位、政治影响力、话语权力、心理情感和精神世界以及文化价值[viii]。经济学不计算这些事情，伦理学却不可不考虑。由于幸福就是生活的目的，排除了幸福问题就等于否定了所有关于生活的研究，特别是伦理学，因此我首推幸福原则；（2）公正原则。在《论可能生活》里，我有比较复杂的表述和论证，在此不论。但它大致可以表达为金规则改进版“人所不欲勿施于人”。

## 7. 遗留问题

由于我的金规则改进版仍然只是关于公正原理的一个表述，它当然就不能覆盖伦理学的所有问题。但是假如把另一条基本原则考虑在内，或许就能够解释大多数问题。所以，金规则至少需要两条，而不是传统格式的一条。但为什么说两条原则（幸福原则和公正原则）的体系只是解决大多数问题而不是全部问题？借助这样一个想象也许可以发现某种潜在的困难：我们可以增加“贿赂问题”的挑战力度，把王庆节设想的“腐败官员共同体”夸张为“所有人”（至少是“大多数人”）这样一个大规模集合，出现这种人类性的集体堕落并非不可能，至少在逻辑上是可能的，这样怎么办？显然，假如一种堕落变成人人同意的，那就不可能出现关于这种事情的任何反面评价了。这种“众心一致”的难题非常值得深思（无论是异口同声说假话还是万众一心做坏事），它有可能造成无法逆转的颠倒黑白。

人类全体堕落的例子也许没有，但人类大体堕落的例子已经不少，现代社会就最善于领导人民走向大体堕落。现代商业和工业所领导的就是一场人类集体堕落的文化运动，它以大众“喜闻乐见”为理由把生活/文化的结构由“向高看齐”颠覆成“向低看齐”，从而消除了文化和精神的品级制度，以低俗的文化精神淘汰高贵的文化精神。在这个关于文化价值的伦理学问题上恐怕很难有普遍原则，这就形成了伦理道德的一个永远的溃口。

[i] 我曾经提出这个“接受问题”去批评哈贝马斯。参见赵汀阳：Understanding & Acceptance. In Les Assises de la Connaissance Reciproque, Le Robert, Paris, 2003.

[ii] Levinas: Ethics as First Philosophy, in The Levinas Reader, ed. S. Hand, Blackwell, 1989.

[iii] Carl Schmitt: The Concept of Political. The Univ. of Chicago Pr., 1996.

[iv] 孔汉思/库舍尔：《全球伦理：世界宗教议会宣言》。四川人民出版社，1997。

[v] Kant: Groundwork of Metaphysics of Morals. Harper & Row, New York, 1964. Section II.

[vi] cf. Henry Sidgwick: The Methods of Ethics. Macmillan, 1890, p.380.

[vii] 王庆节：《解释学、海德格尔与儒道今释》。见“道德金律与普世伦理的可能性”。中国人民大学出版社，2004。

[viii] 一般情况是，看到别人经济上比较成功，有些人会嫉妒而痛苦，这样虽然不好，却不算变态。也许有的人看到别人幸福也感到痛苦，这就是变态了，因为背后是完全扭曲的、反对任何人的极端心理。这些变态的情况是心理学的课题，伦理学不考虑在内，所以不能构成反例。

《中国社会科学》提要：道德金规则是伦理学中公正原理的一个浓缩表达。本文分析了传统的道德金规则的典型民间版本和学术版本，指出传统金规则要得以成立，在逻辑上至少需要两个条件：人际共识和价值共识。而在今天，价值共识已经不存在，所以传统金规则不再成立。作者试图通过从“主体观点”到“他者观点”的转换而提出一个关于道德金规则的新版本，即以“人所不欲，勿施于人”代替“己所不欲，勿施于人”，并试图论证其逻辑上的完美性。本文中的核心论点最早于2000年提出，由于该问题涉及到生活基本准则，事关重大，因此经过听取多方意见和反思，在这里进行了更严格的论证和修改。

## 1. 问题背景

首先需要提及为什么要讨论道德金规则的问题背景。

一般地说，金规则指的是能够概括地表达一个伦理体系的总精神的一条道德原则，也就是能够“一以贯之”的普遍原则；如果在学术意义上说，它就是伦理体系的一个元定理，它是对伦理体系中各种具体规则的总指导和解释。金规则总是非常稳定的，除非社会发生巨大变化，否则它不会变化，而现在正是一个巨变的时代。

全球化把以前不明显的许多问题变成了明显的问题，其中一个典型问题就是各种文化/知识体系之间的关系。亨廷顿关于文明冲突的论点虽然有着根本性的错误，但它却揭开了“对话/交往”问题的真正底牌。自苏格拉底以来，理性对话就被认为是通向普遍承认的真理之路，到今天，哈贝马斯还坚持认为，完全合乎理性标准的正确对话必定能够产生一致认可的理解。但是哈贝马斯忽略了一个关键性的问题底牌，这就是，理解不能保证接受[i]。理性对话有可能达到一致的理解，但是人们想要的不仅仅是被理解，而是被接受。接受才是“对话/交往”问题的终点，这一危险的底牌在以前的哲学分析中被有意无意地回避了。达成共识和充分的理由不是互相理解，而是互相接受。可是互相接受的问题超出了知识论和理性所能够处理的范围。显然，“接受问题”迫使知识论上的“主体间”问题深化为实践或价值理论上的“人际”问题，如果扩大计算单位，则成为“文化间”问题。人们在考虑知识时必定同时考虑价值，知识问题和价值问题是共轭的。这样就回到了哲学的正宗模式上了，无论希腊还是中国先秦，知识问题都是从属于伦理/政治问题的。

当把“接受问题”计算在内，“对话”就变成了“对待”，伦理学就成为第一哲学的一个部分（列维纳斯甚至相信第一哲学只能是伦理学[ii]），另一个部分非常可能是政治哲学（施米特相信政治生活是最基本的生活形式，而政治问题就是区分敌友[iii]）。伦理学和政治哲学的一个最基本问题就是“如何对待他人”。列维纳斯非常正确地论证了他人的绝对性，他人是一个无论如何无法被“我”的主观性所消化的外在绝对存在，主观性（subjectivity）化不掉他者性（otherness），所以他人超越了我的主观性，是我的生存条件和外在环境。他人会反抗，他可以不合作，所以超越了我，所以他人是我需要对待的最严肃的问题。今天人们特别感兴趣的全球合作、全球共识、全球价值之类，在学理上都依赖着关于“他人”的理论。对待他人的总原则在伦理体系中就表现为所谓的“金规则”。

各种文化的伦理体系中都有着至少一个被认为在理论上无懈可击、在实践上历久常新因此非常可能是万世不移的道

德原则，它被当作是一个伦理体系的基石。通常人们把这类据信为“绝对无疑的”道德原则按照基督教伦理的惯例称为“金规则”（the golden rule）。根据孔汉思和库舍尔的研究[iv]，据说不仅各种文化中都有金规则，而且这些在历史中各自独立地自发生成并且以不同方式表述出来的金规则在含义上“都惊人地相似”，几乎可以说其逻辑语义是完全一致的。于是他们认为，这种一致性表明了金规则是放之四海而皆准的普遍必然原则。

可是由经验巧合去推论普遍必然性，这在理论上说（根据休谟定理），是不可以接受的，因为不存在这样一种逻辑，或者说不存在这样一个必然有效的推论模式。在理论上，道德金规则一直没有被成功地证明。而在实践上，虽然金规则一直在伦理体系中有着最重要的意义，但也一直存在着一些根本性的困难，这些困难在古代社会里也许不很明显，可是在当代社会里就变成了严重挑战，比如说，在多文化共享的社会空间里，各种文化（各种文明、宗教、传统和政治理想）甚至各种亚文化（女性主义、环保主义、同性共同体等）都拥有不同的价值观，这意味着它们在“想要的和不想要的”东西上有着不可通约的需要和评价标准，“己所不欲勿施于人”这样的金规则在这里已经没有能力处理那些价值问题了，至少可以说，在许多事情上都会遇到严重挑战。在这个意义上，基于目前被普遍承认的金规则的所谓“全球伦理”是不可能成功的，因为这一全球伦理运动仅仅考虑到金规则在“空间”中的普遍存在，而没有考虑到社会关系在“时间”中的变化。从空间的角度去看，人们似乎都承认有那么一种古老而普遍的金规则，但是，问题已经在时间中发生了巨变，问题变了，答案自然也应该有所变化，传统金规则格式不再是无懈可击的了。

## 2. 人际共识和价值共识

如上所述，金规则表达的是如何对待他人的人际共识。传统金规则有许多版本（在后面我们再作分析），但无论什么版本，其根本精神是完全一致的，或者说，它们的逻辑语义是同样的。传统金规则的基本假定是所有人（或至少大多数人）具有价值共识，也就是所谓“人同此心，心同此理”。与这个基本假定相配合，其方法论则是“推己及人”。以此假定和方法论就必然得出传统金规则。

问题正在于此：诸如“己所不欲勿施于人”这样的人际共识是有效的，当且仅当（iff），一个社会具有共同价值观，也就是具有关于“想要的/不想要的”的价值共识。这两个共识必须同时存在，否则传统金规则就不可能成立。为什么古典的金规则在今天遇到困难了？其秘密就在于现代社会失去了价值共识，因此原来的人际共识就失去了得以成立的必要条件。

这个问题在古代社会所以没有出现，是因为任何一个古代社会都还没有发展出许多互相冲突又几乎同样有力的价值观，即使人们在价值问题上有某些不同意见，也还没有形成各种同样有影响的权力话语，就是说，在古代社会，人们想要的和不想要的基本一致或者说大同小异，即使有些另类人物的奇谈怪论，也只是一些学术性的观点而没有成为有社会影响力的话语，没有成为主流，因此不影响社会的总体价值选择。比较粗略地说，古代社会的冲突的主要原因是利益问题而不是价值观问题。尽管古代也有价值观冲突，但基本上只是学术现象。价值观冲突成为社会现象是现代社会的产物，是所谓启蒙的产物，是自由和平等的结果。很容易想象，在价值观基本一致的社会条件下，就有了“人同此心，心同此理”的普遍现象，于是“推己及人”的方法论就能够适用，金规则的古典版本就是顺理成章的了。可是今天不再如此。

当然，必须承认，即使在价值观基本一致的古代社会里，例外情况总是有的。但是，更应该强调的是，例外的社会现象对于社会一般知识不能构成挑战。这是个有趣的知识论问题，可以称作“例外”的知识论问题。对于自然科学尤其是逻辑知识来说，如果万一出现了“例外”，就是无比严重的问题，因为“例外”构成了挑战普遍必然性的“反例”。可是对于社会生活来说，“例外”是软弱无力的，因为“例外”无法构成对主流价值或者统治性话语的挑战，往往可以忽略不计，“反例”对于社会知识来说是非常可笑的。在古代社会里虽然总有某些例外的价值观点，但被主流价值观所淹没。所以可以说，古代社会具有价值共识。

作为传统金规则的必要条件的价值共识是在现代被破坏的。事实上，从现代开始以来，传统金规则就注定要出现问题了。从价值观念方面去看，现代开始于“平民”反对“贵族”的价值观以及相应的制度安排，自由和平等的要求注定了价值观的多元化和冲突。现代社会的产生当然有着各种各样的重要原因，不过其中的价值观革命可能是最深刻的。尼采可能最早意识到现代性意味着一种彻底革命的价值观，他指出现代就是奴隶反对主人的运动，当然也就要用属于奴隶的“低贱的”价值观去反对主人的“高贵的”价值观。后来，列奥·斯特劳斯又指出现代性的另一个相关的基本精神是“青年反对老年”，也就是今天反对古代——青年被用作隐喻指示现代，因此，他认为“古今之争”是最大的价值观冲突。“青年”这一隐喻意味深长，它不仅可以说现代以“进步/落后”的技术指标替代了传统的“好/坏”人性标准，以“新/旧”的时尚指标替代了传统的“卓越/拙劣”的品质标准，而且还因此导致了无法止步永不停息的“推陈出新”运动，从积极的方面看，这是过不完的青春期，从消极的方面看，这又使得精神积累不再可能。这就是现代性。

在这里我们关心的不是关于现代性的批判，而是想说，现代性这种“新/旧”和“进步/落后”的价值指标必然形成各种各样价值观的大量生产和互相冲突，因为各种价值观都有理由以“新”和“进步”作为其合法性根据。这样来



看，现代性就是没有一致价值观的时代。人的解放导致思想解放，思想解放导致价值多元。正如前面所分析的，一旦失去“价值共识”这个基础，传统金规则就会失去普遍有效性。其实不仅仅是金规则，几乎所有古典的标准都因为现代社会的各种新价值（平等、个人主义、进步、新奇、数量化、多元化等等）而失去效力。金规则问题可以看作是现代问题的一个典型案例，它表明在这个彻底现代化的时代，各种原来认为的普遍原则也都不得不更新换代。

### 3. 对等性结构和互换性结构

道德金规则，无论什么样的版本，都意味着伦理体系的一个元定理。我们知道，任何一个规则系统（哥德尔意义上的严格系统），如果其内容足够丰富的话，就必定是不完备的。而伦理系统本来就是而且只能是很不严格的系统，而且其系统内容必定极其丰富，因此漏洞百出就不足为奇。按照我在《论可能生活》里的论证，一条伦理规则要应用的情景几乎是无穷多的，而生活情景不可能完全一样，因此任何一条规则总是不得不根据具体情景被灵活解释。这种规则与实践的差距就难免导致“标准失控”的难题。于是，要维持伦理系统的解释在大体上的稳定性（绝对严格的稳定性是不可能的），就必须有一些明显普遍有效的一般理念来对各种具体规范作出最后的解释和判断。也就是说，生活的可能情景无穷多，而且其变化情况无法预料，伦理规范本身又总是含糊的（例如“不许说谎”就是一个含糊表达，而如果说成“无论何时何地都对任何人都决不说谎”就会遇到严重困难），因此需要有某些能够应付“所有情况”的基本理念，也就是能够“以不变应万变”的原则来给出最后解释。这些基本理念就是幸福、公正、自由的理念（有的伦理体系还要求更多的基本理念）。对这些根本性的理念的表述可以是一些复杂的理论原则，同时往往也表述为清楚明白的实践原则，这些原则就成为伦理体系的元定理，也就是有能力对某个规则体系的总体性质进行反思并且作出判断的“最后”定理。

通常所谓的金规则正是表达着公正的理念，因为无论在理论上还是实践上，公正原则都是对“如何对待他人”这一问题的唯一理性回答。公正原则对任何涉及他人的行为规范做出理性的判断和解释，因此意味着能够普遍承认的人际关系原则。

在传统金规则的各种版本中，最有名的也可能是最典型的是基督教的金规则和孔子原则。基督教金规则的正面表述是：“你若愿意别人对你这样做，你就应当对别人也这样做”；其反面表述则是：“你若不愿意别人对你这样做，你就不应当对别人这样做”。孔子的正面说法是：“己欲立而立人，己欲达而达人”，其反面说法是：“己所不欲，勿施于人”。

学院派哲学家对这些民间风格的表述不太满意，于是又有了一些学术版的表述，最著名的是康德版，即他的道德普遍律令：“你只能按照你希望能够成为普遍规律的行为准则去行为”[v]。康德认为只有他给出的这个伦理原则才是真正严格的，因为它不需要利用实践经验，仅仅通过理性本身而获得的。基于民间经验的金规则被认为应该以基于理性本身的原则为准去重新理解。还有比康德版更细致一些的西季维克版（但未必比康德版高明）。西季维克指出，金规则的正面表述肯定是错误的，因为人们完全可能愿意互相帮助做坏事。金规则的反面表述虽然不是错的，但却仍然不准确，西季维克的修改版大概是这样的：“对于任意两个不同的个体，A与B，如果他们各自情形上的不同并不足以成为在道德上加以区别对待的根据，那么，如果A对B的行为不能反过来同时使B对A的同样行为同样是正确的话，这一行为就不能被称为在道德上正确的行为”[vi]。如此等等。尽管哲学家们相信他们的学术版比民间版严格得多，但仍存在着严重问题。真正根本的问题并不在于什么样的表述更为严格，而在于所有传统版本，无论是民间版还是学术版，都有着同样错误的思想出发点。学术版尽管严格，却只不过严格地继承了民间版的错误。其中比较好的版本应该是孔子的正面表述，尽管正面表述的金规则比较冒险，容易出现明显的漏洞，但孔子的正面表述包含了“立”和“达”这样的模糊概念，就多少回避了困难，但基本的困难仍然存在。

不管那些学术版的金规则如何改进表述上的逻辑性和严格性，但各种金规则的基本结构仍然是一致的，即强调互相对待的对等性（reciprocity）。这个结构没有错误。对等性结构应该是表述“我与他人”的关系的最合理结构，因为必须有对等性结构才能够表达公正关系，而金规则就是为了表达公正关系的。在《论可能生活》（1994年版；2004年版）中，我也是以“对等性”为基本结构来分析和定义公正的（在定义公正时，一部分哲学家主要强调“对等性”，另一些哲学家则主要强调“公平性”，这是两个众望所归的基础，但在本质上说，“对等性”是更为重要的因素，只有当承认了对等性，公平性才有意义。对等性固然不错，但事情没有那么简单，我们还需要警惕另外某些可能被漏掉的因素。

公正在形式意义上直接就意味着一种对等性。人们早就意识到公正是一个“恰如其份”的概念，它意味着各得其所、各得所值。无论对于人际关系还是事际关系，公正的对等性首先表现为“等价交换原则”，即某人以某种方式对待他人，所以他人也以这种方式对他，或者某人以某种东西与他人交换与之等值的東西。这一原则虽然非常“清楚明白”，就像笛卡儿所推崇的真理那样，但它实际有效的情况却很有限，因为，只有当双方在某种情境中具有几乎同等的自由和能力时，这一原则才能够被有效地执行。于是，公正的对等性通常又表现为一些比较复杂的对等形式，比如“礼让原则”，即某人以对待什么人的方式对待我，那么我就以什么人的方式回报他（如礼让所说“……以国士遇臣，

臣故国士报之”)。还有“西季维克原则”:给同样的事情以同样的待遇,而给不同的事情以不同的待遇。

然而,公正仅仅表现为对等性是不够的。怎样才真正算是对等的?这仍然是不够清楚的事情。对等原则并不能解决需要公正处理的所有问题。具体地说,等价交换原则,即A以X方式对B,因此B有理由以X方式对A,只能证明“B以X方式对A”是公正的,却无法证明“A以X方式对B”是公正的;而礼让原则,即A以(B=X)的方式对B,因此B有理由以X的方式对A,也只能证明“B以X的方式对A”的公正性,却不能证明(B=X)这一方式的公正性;同样,西季维克原则,可以表述为“按照标准X,A和B是同样的,所以给予A和B同样待遇”,也只能说明对于给定标准X,A和B得到同等待遇是公正的,却无法证明设定标准X是不是公正的,我们也就无法真正知道A和B被看成是同样的是否是公正的。这些情况的共同问题表明,在对等性结构中,我们只能必然地证明“后发行为”的公正性,却没有理由证明“始发行为”的公正性。这是传统公正理论以及传统金规则所难以处理的问题。

为了使一个相互关系得到在场各方的共同绝对认可——排除了迫于条件、压力和强迫的相对认可——就必须使得在场各方都认可这样一个“地位互换”原则:如果A以X方式对待B是正当的,当且仅当,A认可“当A处于B的地位而B处于A的地位,并且B以X方式对待A是正当的”。对于这种地位互换关系,如果无论把“我”代入为A或B,“我”都将认可其中的行为方式,那么这一行为方式就是正当的。这个地位互换原则在利益分配上同样有效,它表现为:如果A按照X准则把A和B看成是同等的,并且A和B得到同等的利益分配是正当的,当且仅当,A认可“当A处于B的位置而B处于A的位置并且B按照X准则把A和B看成是同等的,并且因此得到同等利益分配是正当的”。这个“地位互换”原则可以看作是传统的角色互换原则的改进版。传统的角色互换主要表达的是“将心比心”或者“同情”的直观,这个直观大体不错,但是不够严格,因为无论是“我”还是“他人”都似乎是抽象的同样的人,其中暗含着“既然都是人,那么就都应该什么什么”这样的认识,这其实仍然没有超出对等性原则。事实上任何一个社会都存在着不同地位的关系,所以很难简单对等。地位互换原则就是试图发现当把地位差异考虑在内时能够产生什么样的公正关系。

现在我们可以比较完整地理解公正的意义了。首先,公正表现为对等性。这意味着允许存在着某种假设X,然后在X的基础上要求对等。这样可以保证在给定价值标准下的公正关系;其次,公正进一步表现为互换性。这意味着任何一个可能的假设X,即使它能保证对等,也必须被证明为在互换方式中是有效的。这一地位互换原则可以消除在设立价值评价标准上的不公正。

#### 4. 想象一个最佳版本

我曾经提出对金规则进行过一个本质性的修改,当然没有藐视传统智慧的意思,而仅仅是在新的问题框架中发展了传统智慧。既然金规则的反面表述被普遍认为是比较稳妥的格式,那么,我选取孔子的反面表述“己所不欲勿施于人”作为底本(包括基督教金规则在内的各种反面表述版本的语义都可以完全表达在孔子版本中),把它修改为“人所不欲勿施于人”。仅仅一字之差,但其中的人际关系发生根本性的变化。我认为这是目前所能够想象的最佳版本,它至少具有两个在理论上或者说在技术上的优势:1)假定“人所不欲勿施于人”成为替代“己所不欲勿施于人”的一种新的人际共识,它不需要以价值共识作为必要条件,而既然免除了价值共识这个苛刻要求,它因此就可以良好地适用于今天社会的价值多元情况,而且是克服由于价值多元而产生的文化冲突的一个有效原则;2)尤其在纯粹理论上,“人所不欲勿施于人”原则能够满足严格意义上的普遍有效性要求,相比之下,“己所不欲勿施于人”原则只是在特定条件下有效的,并非真正的普遍原则。实际上,传统金规则将成为新版金规则所蕴涵的一个特例,就是说,假如一个社会碰巧具有价值共识,在这个特殊条件下,“人所不欲勿施于人”与“己所不欲勿施于人”意义等值,但是在其他社会条件中,就只能是“人所不欲勿施于人”。就是说,“人所不欲勿施于人”的有效范围大于“己所不欲勿施于人”。

#### 5. “无人被排挤”原则

现在进一步来说明把“己所不欲勿施于人”修改成“人所不欲勿施于人”的更深入的哲学理由。

从思想语法上看,人们在思考“我与他人”的关系时一直使用的是主体观点,即以“我”(或特定统一群体“我们”)作为中心,作为“眼睛”,作为决定者,试图以“我”为准,由“我”来定义知识和标准,按照“我”的知识、话语、规则把“与我异者”组织为、理解为、归化为“与我同者”。这种自我中心的态度可能是一种自然态度,恐怕与文化无关,即使是从来都非常强调他人重要性的中国传统文化,显然也没有完全超越自我中心的理解方式。在列维纳斯看来,这种传统的主体观点是对他人的不公。一切以我的观点为准,实际上就是对他人的否定,是实施了一种无形的暴力,一种试图化他为自己的暴力。把这种态度作为我与他人关系的道德基础显然是无效的。我们必须以把他人尊称为“您”的他人观点来代替传统的主体观点,只有以他人观点为准的理解才能尊重他人存在的“超越性”,即不会被“我”随便“化”掉的绝对性,才能避免把他人的超越性消灭在我的“万物一体化”的企图。尊重他人的

超越性就是尊重他人的那种不能被规划、不能被封闭起来的无限性。

列维纳斯的观点或许有些夸张，但确实指出了问题之所在。我们无须把他人夸张为至高至尊的存在，但至少可以发现这样一个逻辑关系：尊重他人的同时就等于让自己得到尊重，因为对于别的主体来说，自己也是个必须被尊重的超越者。只有从他人观点出发才能演绎出真正公正的互相关系，即列维纳斯所谓的由“面对面”所形成的“我与你”关系。“我与你”的关系是平等的，而“我与他人”的关系则是不平等的，按照康德的说法，它是把人当作了“手段”的关系。列维纳斯这一“他人为尊”观点在思想语法上非常重要。如果从主体观点入手，心中始终以我为准，我把你贬低为在我的知识和权力范围里去解释的他人，你当然也会把我当成他人，这种似乎对等的关系逻辑地导致了实际上的互相否定，这是一种互相拆台的对应。即使好象有了互相尊重的良好意愿，但只要思维方式是主体观点，那种意愿就无从实现（康德就希望仅仅以“善良意志”作为唯一条件，好象善良意志足够解决问题。可是良好意愿有时只不过是欺骗自己说自己是个好人，对他人并无实质好处，如果仅仅有意愿而没有行动，只能证明意愿的落空）。

根据列维纳斯的批判，我们可以有进一步的发现：所谓各种文化都认可的、据说可以作为第一普遍原则的金规则“己所不欲勿施于人”表面上似乎表达了对他人的善意，其实隐藏着非常典型的主体观点霸权，它的思维出发点仍然是“己”，它只考虑到我不想要的东西就不要强加于人，根本没有去想他人真正想要的是什么，这意味着，我才有权利判断什么东西是普遍可欲和什么事情才是应该做的，我的心灵才算是个有资格做决定的心灵，而他人的心灵和思想根本不需要在场，我可以单方面决定普遍的价值选择。这就是主体性的霸权。

取消他人的价值决定权是非常严重的错误，这里存在着深刻的权利之争，它是知识权利之争，它是关于知识的政治斗争，即谁才有权利决定什么算是合法的或正当的知识，或者说，谁才有权利决定什么算是有价值的东西以及什么算是正当的事情。知识权利不仅是政治权利，而且是最根本的政治权利。所以，即使像“己所不欲勿施于人”这样的金规则的确表达了善良意志，这种心理学的解决方式也没有能够真正解决我与他人的公正关系问题。可以说，伦理学问题的解决不能依靠心理学，无论是康德式的善良意志还是孟子式的良心，而只能指望一种政治学的解决，即必须能够使得我与他人同样拥有知识上的政治权利。

可以进一步分析，为什么主体观点在伦理学中是不合法的？我们知道，主体观点首先是一种知识论观点，是关于如何“看”世界的方法。在“看”事物时，不需要征求事物的同意，我想怎么看就怎么看。在纯粹知识论里，主体观点虽然不见得是最优的（它所欠缺的视角太多），但却是合法的。但是在其它地方，主体观点就失去了合法性。例如在交往问题中，我就不可以想怎么就怎么随便“看”别人，因为问题已经由“看”变成了“说”和“听”，其中关键还落在“听”上，我们不得不去听别人说什么，否则交往不成立，在这里，主体观点的合法性就已经很可疑了。当进入到“做”的问题，即共处或合作的问题（做任何事情都涉及他人），主体观点就完全不合法了，因为生活问题的基础是存在论而不是知识论，把知识论的原则推广到存在论中去是不合法的，显然，他人不可以被处理成像物那样的“对象”，因为他人意味着我所化不掉的外在精神，他也将按照他的精神去做事情，而且我要做任何事情都不得与他人一起做（至少是间接地与他人一起做事情）。因此，在存在论所控制的领域里，他者性（otherness）就必定成为替换主体性（subjectivity）的最高原则。

当把他人观点计算在内，道德原则所考虑到的变量就有了巨大的变化，既然增加了必须计算在内的变量，我就有理由认为应该把金规则修改为“人所不欲勿施于人”。虽然只一字之差，但其中境界却天上地下。在“由己及人”的方法论中，可能眼界只有一个，即“我”的眼界，而“由人至人”的方法论则包含了所有的可能眼界（all possible horizons），这样才有可能尊重每个人。这个把每个人的精神考虑在内的方法也可以称为“无人被排挤”的原则。这一改变非常必要，它可以根本改变我们思考价值问题的思路和角度。在这里它意味着一条关于选择任何共同伦理的元规则，这个元规则蕴涵着彻底的公正，可以表述为：（1）以你同意的方式对待你，当且仅当，你以我同意的方式对待我；（2）任何一种文化都有建立自己的文化目标、生活目的和价值系统的权利，即建立自己的关于优越性（virtue）概念的权利，并且，如果文化间存在分歧，则以（1）为准。这是学术版的“人所不欲勿施于人”。

## 6. 进一步的论辩

对于我的金规则修改版（我曾经在先前的某些论文中简单地提及这个观点），有一些不同意见。目前最具学术深度的挑战是王庆节先生提出的。王庆节指出：“赵汀阳以‘他人观点’建构起来的道德金律固然可能帮助我们克服‘主体观点’，但这种以‘他人观点’为准的道德金律是否能胜任作为‘普世伦理’的‘元规则’呢？我想我们恐怕不能得出这一结论。为什么呢？让我们来看下面的例子。假设你我都是腐败的官员，而且你我都不以贿赂为耻，反而以之为荣。当我贿赂你时，我知道你想我以“贿赂”的方式对待你，并且假设你也会同意以同样的方式回报我。但是，我们知道，按这种方式进行的行为，无论是出于“主体观点”还是“他人观点”，都不能改变“贿赂”的不道德性。这也就是说，即便我对某一他人对我行为的所欲所求与此人想要得到的对待是相同的，遵循这一原则行事也不能保证永远是道德的” [vii]。

这一批评的挑战是富有启发性的，而且在某个方面来看是正确的。它试图指出任何可能设想的形式化道德原则的弱点。类似的例子曾经被用来批评康德的形式化的“普遍律令”。这是一个值得深思的问题。这种挑战与分析哲学家们喜欢的“反例”完全不同，它不是一个具体反例而是一种逻辑可能性，这样就具有力量。我们前面提到“反例知识论”问题，我相信个别反例对于社会科学知识没有挑战意义，因为社会科学所要求的知识标准不是像科学和数学那样的“无一例外”标准，它关心的是具有社会影响力的存在，个别反例无法改变任何主流的社会事实，所以个别反例对于社会科学几乎没有意义。但是，逻辑可能性则是现实的威胁，这不能不考虑。在深入分析王庆节的问题之前，可以首先考虑一个技术性问题以便削弱其挑战力度。王庆节的关于“贿赂”的假想可能性对于批评金规则的正面表述比较有效，而对于反面表述，情况则有些不同。我们所讨论的是人所“不欲”，某人A想要贿赂，这显然属于“欲”而非“不欲”，那么，是否可以把命题“A想要贿赂”兑换成命题“A不想清廉”？或许这两个命题在真值上是逻辑等值的，但却并非意义等值。显然，受贿是个经过决定而做出的行为，而“清廉”是个状态，是个还没有发生受贿事件的无为状态，一个人在没打算受贿的时候，他的本来状态就是清廉的，所以，只有“受贿”才是个欲望对象，我们似乎不能把初始状态看成欲望对象。这有点像希腊诡辩论者的例子，显然不能由我们“本来没有角”推论说我们“失去了角”。所以，“贿赂案例”只能构成对正面表述的金规则的挑战，却不太可能对反面表述的新版本金规则形成真正的威胁。

虽然以上分析削弱了“贿赂案例”的挑战性，但我们仍然必须承认，总会有办法找到某些具有挑战性的方案，不过恐怕不可能对改进了的金规则形成釜底抽薪式的威胁，但是会迫使人们承认金规则的局限性。比如说，有人喜欢战争、侵略和压迫，而和平、友好和尊重正是他所“不欲”的；又有人愿意吸毒、赌博加性变态，而改邪归正是他所“不欲”的，在这里，新版本金规则并不能提供足够理由去反对他。其实，无论什么样的金规则都有局限，就是说，任何金规则所能够治理的伦理空间是有限的，总有很大的“伦理荒地”留在外面。

康德非常可能意识到了这个弱点，这也许就是康德为什么除了形式化的“普遍律令”之外还要给出了另一条基本原则的理由。另一条原则就是甚至更为知名的“人是目的”原则。可惜这条原则的要求太高，以至于是不切实际的空话，人们不可能不在某些场合把某些人当成“手段”，否则生活不可能进行，更不可能有什么社会了。正如我在《论可能生活》里论证过的，除了单纯美好的人际关系之外，人际关系在许多必要的情况下表现为各为其利的“事际关系”，所谓“手段”的问题就必然出现，如果不把某些人当成手段，就根本做不成任何事情。不过重要的是，“人是目的”原则不再是纯形式的了，而是有价值内容的（主要是人道主义的价值观）。这样就指出了在金规则范围之外的伦理问题，也就是说，伦理学的基本问题不仅仅是公正问题，而且还有价值观的选择问题；不仅是为了解决我与他人的关系问题，而且还需要解决生活和社会的理想问题。尽管康德所选择的基本原则很有疑问，但他构造的道德原则框架确实不同凡响，难怪他可以声称他的道德原理比金规则要好得多，而且金规则必须按照他的原则进行校正。且不管人类道德体系所最需要的内容原则是什么，至少我们可以肯定，仅仅表现为形式原则的金规则是不充分的，显然还需要内容方面的约束条件。就是说，金规则不可能只是一条，而是至少需要两条，而且这两条原则是互相约束着的。

在某种意义上，康德方案既不现实又有自相矛盾之处。假定“人是目的”原则普遍成立，那么就会有一个完美到不可能出现和存在的社会。假如真有完美社会，现在我们观察到的绝大多数社会规律和规则，以及绝大多数的人文社会科学知识就失去意义，尤其是经济学、社会学和政治学的全部前提就不存在了。如果“人的目的”，人人自然就能够公正地互相对待，那条用来表达公正的形式原则就是多余的了。而既然需要公正原则，就是因为人们只能指望策略性公正，而不能指望无论何时何地都“人人成目的”的绝对公正。因此，我们有理由认为，康德的框架是伟大的，但其内在构造并不妥当，两条原则并不能很好匹配。要选择两条良好匹配的、同时又足够普遍以至于能够覆盖各种伦理学问题的“道德基本原则”确实不容易。在《论可能生活》中，我采取的也是“两条原则”的模式，但与康德所选择的原則有所不同，解释也很不同，尤其是与康德的道德原则排序不同，康德首推形式原则，然后才是价值原则，我的排序正好相反。

目前我的方案是（希望以后能够改进）：（1）幸福原则：如果一个人的某个行动是自成目的的（autotelic），那么这是一个必然产生个人幸福的行动，而如果连同这个行动的目标也是一个自成目的的事情，那么就同时成为一个能够促进人类幸福的行动。幸福既是人们的最大欲望，又是唯一对己对人都有益的行动。“幸福的帕雷托改进”是唯一具有全方位正面效果的帕雷托改进，就是说，某人的幸福不仅对自己有益，而且在任何意义上都对他人无损。相比之下，福利的帕雷托改进（即经济学意义上的帕雷托改进）虽然无损于任何他人的物质利益，但却非常可能损害他人的社会地位、政治影响力、话语权力、心理情感和精神世界以及文化价值[viii]。经济学不计算这些事情，伦理学却不可不考虑。由于幸福就是生活的目的，排除了幸福问题就等于否定了所有关于生活的研究，特别是伦理学，因此我首推幸福原则；（2）公正原则。在《论可能生活》里，我有比较复杂的表述和论证，在此不论。但它大致可以表达为金规则改进版“人所不欲勿施于人”。

## 7. 遗留问题

由于我的金规则改进版仍然只是关于公正原理的一个表述，它当然就不能覆盖伦理学的所有问题。但是假如把另一条基本原则考虑在内，或许就能够解释大多数问题。所以，金规则至少需要两条，而不是传统格式的一条。但为什么说两条原则（幸福原则和公正原则）的体系只是解决大多数问题而不是全部问题？借助这样一个想象也许可以发现某种潜在的困难：我们可以增加“贿赂问题”的挑战力度，把王庆节设想的“腐败官员共同体”夸张为“所有人”（至少是“大多数人”）这样一个大规模集合，出现这种人类性的集体堕落并非不可能，至少在逻辑上是可能的，这样怎么办？显然，假如一种堕落变成人人同意的，那就不可能出现关于这种事情的任何反面评价了。这种“众心一致”的难题非常值得深思（无论是异口同声说假话还是万众一心做坏事），它有可能造成无法逆转的颠倒黑白。（<http://www.tech.cn>）

人类全体堕落的例子也许没有，但人类大体堕落的例子已经不少，现代社会就最善于领导人民走向大体堕落。现代商业和工业所领导的就是一场人类集体堕落的文化运动，它以大众“喜闻乐见”为理由把生活/文化的结构由“向高看齐”颠覆成“向低看齐”，从而消除了文化和精神的品级制度，以低俗的文化精神淘汰高贵的文化精神。在这个关于文化价值的伦理学问题上恐怕很难有普遍原则，这就形成了伦理道德的一个永远的溃口。（<http://www.tech.cn>）

[i] 我曾经提出这个“接受问题”去批评哈贝马斯。参见赵汀阳：Understanding & Acceptance. In *Les Assises de la Connaissance Reciproque*, Le Robert, Paris, 2003.（<http://www.tech.cn>）

[ii] Levinas: *Ethics as First Philosophy*, in *The Levinas Reader*, ed. S. Hand, Blackwell, 1989.

[iii] Carl Schmitt: *The Concept of Political*. The Univ. of Chicago Pr., 1996.

[iv] 孔汉思/库舍尔：《全球伦理：世界宗教议会宣言》。四川人民出版社，1997。

[v] Kant: *Groundwork of Metaphysics of Morals*. Harper & Row, New York, 1964. Section II.

[vi] cf. Henry Sidgwick: *The Methods of Ethics*. Macmillan, 1890, p.380.

[vii] 王庆节：《解释学、海德格尔与儒道今释》。见“道德金律与普世伦理的可能性”。中国人民大学出版社，2004。（<http://www.tech.cn>）

[viii] 一般情况是，看到别人经济上比较成功，有些人会嫉妒而痛苦，这样虽然不好，却不算变态。也许有的人看到别人幸福也感到痛苦，这就是变态了，因为背后是完全扭曲的、反对任何人的极端心理。这些变态的情况是心理学的课题，伦理学不考虑在内，所以不能构成反例。（<http://www.tech.cn>）

《中国社会科学》 2005(3)

中国社会科学院应用伦理研究中心  
北京建国门内大街5号 邮政编码：100732 电话与传真：0086-10-85195511  
电子信箱：cassethics@yahoo.com.cn