Copyright@United Nations, 2015. All rights reserved

附 件 6

可理解性测试方法

Copyright@United Nations, 2015. All rights reserved

Copyright@United Nations, 2015. All rights reserved

附件 6

可理解性测试方法 [1]

A6.1　本方法的目的，是为了提供一套方法，用于评估化学品危险标签和安全数据单是否易于理解。制定这一工具时，考虑的重点是解决发展中国家工人和消费者的需要。提出这种方法的主要目的，是提供一种尽可能具有全球适用性的工具，同时亦考虑到不同的文化水平和文化背景。

A6.2　方法概述

A6.2.1　本方法分为多个不同的模块，本附件对每个模块的用法作了说明。一般说来，本方法包括以下四个组成部分：

(a) **模块 1**：这是一个核心模块，其主要目的是确保在模块 2 至 11 中所用的方法适合各种不同的文化和背景。建议在所有类别的目标群体中使用这种模块(见下表 A6.2)，但对于文化不同于制做标签和安全数据单背景的工人和社区成员群体来说，必须首先从这种模块开始；

(b) **模块 2 至 8**：这些模块包括普通问卷(模块 2)，和一组关于标签及安全数据单的问题和练习(模块 3 至 8)。根据调查对象的情况，是否是工人，是否使用安全数据单等，这些模块中的某些元素可能不适用；

(c) **模块 9**：这是一种模拟练习。有一个版本对象是工人，适用于生产领域中的绝大多数人；另一个版本(模块 9a)适用于消费者的情况；

(d) **模块 10**：模块 10 包括一份最终的测试后问卷。它适用于问卷调查(模块 2 至 8)和模拟练习(模块 9)的所有参加人。也可用于分组练习的参加人(模块 11)。调查问卷的重点是培训和过去的经验，并提供机会，让人们随意就测试过程提出任何反馈和意见；

(e) **模块 11**：这是供工人使用的分组练习。它吸收了上述模块中的所有元素，目的是在分组学习中测试可理解性。其宗旨是补充模块 2 至模块 10, 但测试的对象与模块 1、模块 2 至模块 8 和模块 9 有所不同。

A6.2.3　此外，还建议在可理解性测试后的一个月和 12 个月，进行后续测试。这种测试应对参加首次测试的人进行重复测试。根据资源和后勤情况，或许可以避免对在底线完成的所有模块重新进行测试。重复测试对于深入了解危险信息接触的保持力和实际益处十分重要。

A6.2.4　表 A6.1 总结了本方法中使用的模块、各模块的主要活动内容，以及每个模块要实现的目的和结果。

A6.2.5　虽然本测试方法是作为独立的成套测试设计的，但根据当地优先事项和需求，从这套测试中选择模块使用也是可以的。而且，人们认识到，危险公示做法的全球统一是一个不断发展的过程，有可能产生新的测试需求。本方法可以调整，随时吸收新的测试优先事项，在相同的测试形式中，使用经过调整的测试材料(标签和安全数据单)。比如，如果正在考虑新的危险符号图标，可修改模块 4, 使之包括新的符号。

A6.3　附件 6 和测试方法的使用

A6.3.1　每个模块都是为一组特定的可理解性测试目的而设计的实际测试问卷。模块的设计是这样的：问卷中清楚地标明了针对进行可理解性测试者的使用说明。伴随着每个模块但分别提出的，是一组

[1] 由开普敦大学一多学科小组为国际劳工局(劳工局)危险公示工作组编写，这是促进全球统一危险公示制度的国际努力的一部分。

Copyright@United Nations, 2015. All rights reserved

详细的指导性说明，它们构成了特定模块的使用手册。使用手册还概述了在每个模块中使用的不同标签和/或安全数据单，以及每个模块的产出和时间要求。

A6.3.2　为避免模块过于冗长，模块使用说明在模块正文中保持在最低限度，而在使用手册中给出详尽的说明。如果模块 3 至模块 11 中有关键说明，它们在阴影框中用黑体字列出，以便于实行。在整个模块中用斜体字标出需要读给测试对象听的所有案文。

A6.3.3　某些模块(模块 3、4、6、7、8 和 9)要求随机选择标签和/或安全数据单。为测试者提供了一盒卡片，以提高一个随机标签/安全数据单或一组标签/安全数据单的选择速度。测试者将获得一个特定卡片盒，上面标有每个有关的模块。

A6.3.4　提供了标签和安全数据单，但它们应符合将要应用这种方法的国家所采用的标准风格和表述方式。全球统一制度将使危险公示方法的内容和设计达到一定程度的标准化，但由于当地的传统习惯、风格、大小和喜好不同，也会出现许多变化。测试中使用的标签和安全数据单必须尽可能反映典型的当地使用习惯。因此，虽然在提供样品标签和安全数据单时附有使用手册，但还是鼓励用户在测试设计要求的范围内对测试材料进行修改，使材料让当地测试对象看上去尽可能可信。

A6.3.5　尽管我们努力简化测量危险公示可理解性所需要的相对比较复杂的测试程序，但测试方法需要谨慎实行，需要进行质量控制。因此，对测试者的培训十分重要。模块 1 和 2 的使用手册对此作了更为详细的说明。

表 A6.1：可理解性测试：按模块分列的目的和结果

模块	内容	目的	结果
模块 1	**核心小组**	使研究工具适合特定目标组的背景、语言和文化解释。 识别词汇的特定文化含义。 测试评级、用颜色划定危险的做法，以及对非明显变量的定量估计是否在文化上是可转化的。 在后续模块中使用的测试方法用于测试表面效度，并确定可选方法。 识别在测试条件下因项目使用的文化差异而可能出现的偏见。	对困难词汇提出在文化上具有一致性的解释。 在当地使用适当的颜色。 说明有可能使可理解性测试引起偏见的文化因素。 确认色盲试验方法。 心理测量尺度对非西方大众具有可解释性。 前后关系测试。 获取工人经验的方法。 “虚拟”符号。
模块 2	**一般性测试**	确定将人口统计学和其他数据作为可理解性分析的基础。 阐明某些后续试验所需要的颜色识别能力和视觉敏锐度。 收集对于可理解性评估解释至为关键的有关工作经验的数据。	与研究结果和分析有关的人口统计学和其他数据。 颜色和视觉敏锐度评估。 工作经验在可理解性方面所起的作用。
模块 3	**标签和安全数据单的回忆、阅读和可理解性**	评估受检对象对标签和安全数据单的熟悉程度。 测试受检对象对标签要素的记忆力。 评估察看标签要素时采用的顺序。 测试信号词、颜色、符号和危险性说明的可理解性。 评价标签对受检对象的影响： - 对自身、配偶或子女的危险评级； - 使用、保存和处置化学品的意向。 在有关可理解性问题提出后，评级和报告是否出现变化。 受检对象是否能够正确识别有关安全数据单？ 受检对象是否能够正确识别有关化学品名称、健康危险、物理危险和使用防护服的信息？	识别对标签和安全数据单的熟悉程度。 评估不同标签字体的影响。 确定不易理解的要素术语。 确定最容易理解的说明。 危险性评级和由于看到标签而打算采取的行动。 将有关可理解性的详细问题对受检对象的危险认识的影响作为替代培训。 对霍索恩效应的影响作出评估。 比较对自身的危险评级是否不同于对近亲的危险评级。 确定受检对象是否能够有意义地将从标签得到的数据与有关的安全数据单联系起来。

Copyright@United Nations, 2015. All rights reserved

表 A6.1：可理解性测试：按模块分列的目的和结果

模块	内容	目的	结果
模块 4	**危险评级与可理解性：信号词、颜色和符号**	测试受检对象对下列内容所表示的危险严重程度的相对评级： - 信号词、颜色和符号； - 符号组合和多种符号； - 符号、颜色和信号词的选定组合。 测试对信号词、颜色和符号的理解程度。 测试对信号词、颜色和符号能否引人注意的看法。 测试受检对象对标签的理解是否会影响他们使用、存贮或处置化学品的原有打算。 研究受检对象对标签上为何出现危险要素的看法。	将对信号词、颜色和符号能否表明危险程度，能否单独理解或组合理解各要素作出评价。 对评级的表面效度的质量控制评估。 标签要素引人注意的能力。 将对引人注意的评级最高的标签的下列能力进行研究： 促使受检对象进一步了解信息，特别是健康危险信息。 影响他们所说的安全行事的打算。
模块 5	**带有或不带文字的危险符号的可理解性**	测试受检对象对代表危险级别的符号的理解。 测试受检对象对危险种类概念的理解。 确定增加文字说明是否可以提高人们对若干代表危险种类：生殖、致癌性和致突变性的符号的理解。 确定增加信号词是否提高人们对代表危险种类的符号的理解。	识别正确的危险种类符号的能力。 确定符号表达效果差的危险种类和确定作为危险种类指标表达效果差的符号。 确定解释含混的符号。 表示生殖、致癌性和致突变性危险的符号增加文字说明的效果。 表示危险种类的符号增加信号词的效果。
模块 6	**符号/象形图的大小、布局、背景颜色和边界**	测试改变符号大小、边界和布局的影响。 测试改变象形图中相对于边界的背景颜色和图标大小的影响。	符号大小、边界和布局对下列各项的影响： - 识别化学品名称的能力； - 对危险性的理解； - 记得符号是吸引人们注意符号的代用品； - 记得危险说明作为一种标识，吸引人们注意力危险说明的代用品； - 报告的行为意向； - 阅读顺序； 比较自身的危险评级是否不同于近亲的危险评级。

Copyright@United Nations, 2015. All rights reserved

表 A6.1：可理解性测试：按模块分列的目的和结果

模块	内容	目的	结果
模块 7	**象形图可理解性 — 附加测试(农药)**	测试受检对象识别下列信息的能力： - 化学品名称； - 健康危险； 对受检对象的危险评级作出评估。 测试受检对象对象形图的理解。 对受检对象的阅读顺序作出评估。	对象形图的可理解性：了解、危险性评级、关注、阅读关键信息。 比较自身的危险评级，是否不同于近亲的危险评级。
模块 8	**按数据组织分列的安全数据单的可理解性**	测试受检对象从安全数据单中识别安全信息的能力。 测试受检对象对安全数据单上的危险信息的理解。 对受检对象在安全数据单上阅读了哪些内容，以及受检对象报告的阅读安全数据单各要素的顺序作出评估。 对哪些信息有用、合适和能够理解作出评估。 评估安全数据单上的信息是否与安全行事的意向有关联。 评价安全数据单信息的不同组织方式对上述各项的影响。	从以下几个方面对安全数据单危险信息的可理解性作出评估： (a) 健康危险信息的解释； (b) 信息对他人的可理解性的自我评价； (c) 就受检对象如何向第三方解释危险说明打分； (d) 报告的行为意向；评价这四种理解措施的一致性。 对以不同方式组织安全数据单信息的影响作出评估。 主观评价各个子要素的有用性和适当性，以确定在安全数据单的补充和完善中需要进一步审查的方面。
模块 9	**模拟练习：标签和安全数据单，以及标签上的符号和信号词的使用对安全化学品做法的影响**	对处理化学品的模拟练习中的安全做法作出评估。 对信号词“危险”的存在和/或危险符号(骷髅和交叉骨)的大小是否改进了安全做法作出评价。 确定过去与化学品有关的经验，是否在安全做法以及信号词和符号对安全做法的影响上都发挥着重要的作用。	对在执行任务之前和执行任务过程中观察到的与标签、安全数据单的使用有关的实际行为作出评价。 安全行为包括 PPE 的使用和其他预防性卫生做法。 改变标签要素(带或不带“危险”字样；不同大小的危险符号)和安全数据单布局(醒目的健康危险标题和在规定信息下给出的健康危险数据)产生的影响。 对理解、做法和试验条件之间的关系进行研究。

Copyright@United Nations, 2015. All rights reserved

表 A6.1：可理解性测试：按模块分列的目的和结果

模块	内容	目的	结果
模块 10	**测试后/模拟测试后**	弄清接触化学品的既往史和培训。 测试符号、信号词、颜色和危险说明的简要解释对危险严重程度评级和可理解性的影响。 确定受检对象的化学品信息需求。	将从培训和过去经验获得的变量用于对模块 3 至 9 的回答进行分层分析。 结果将有助于表明，从长远看是否需要对培训进行更为详细的评价。 对化学品信息需求问题的回答，可能对全球统一制度在化学品安全方面所做的努力有所帮助。
模块 11	**分组练习 — 可理解性**	测试在危险信息的学习方面，分组学习与个人学习之间是否存在差异。 测试受检对象作为小组，在回答同一个问题时，与被单独问到时给出的答案是否显著不同。	对小组学习与个人单独学习的效果作出质量控制评估。 小组的回答与个人的单独回答如果存在显著差异，则说明测试方法需要改进。 对将来如何开展作为危险公示要素的培训的意义。

Copyright@United Nations, 2015. All rights reserved

Copyright@United Nations, 2015. All rights reserved

A6.3.6 同意：在使用本方法中的任何一个模块之前，应先得到参加人的知情同意。为此，应向参加人说明练习的目的，以及要求他们遵循的程序。不得强迫参加人参与，应让他们知道，他们有权在任何时候退出。在同意程序中提供的信息非常普通，不会泄露所测试的明确假设。

A6.3.7 同意程序在模块 1(核心小组)、模块 2(测试开始)和模块 10(模拟练习)的开始部分作了概述。不论同一受检者是否完成所有的模块，都应在需要时进行所有这三项同意程序。模拟练习的同意程序在很大程度上是一种解释，旨在通过向受检者讲明练习目的，消除将会产生的明显偏见。

A6.3.8 参加人奖励或补偿政策：参加这项研究的每个答卷人都应为此而得到某种形式的补偿或奖励。应该在参加人同意参加这项试验时告诉他们，在试验研究结束时，将给予他们某种形式的补偿。各国的补偿可能有所差别，这主要取决于文化背景和当地条件。建议补偿形式(根据其他研究)为正餐(午餐)，帽子、杯子、食品(糖、大米、玉米面)、证书等等。具体的适当补偿政策应由应用这一工具的国家制定。

A6.4 样本采集

A6.4.1 目标人群

A6.4.1.1 目标人群在下表 A6.2 中列出。他们大多数是成年工作人员，属于直接或间接使用、销售或管理化学品的典型群体。儿童是另一个重要的潜在受众。然而，虽然向儿童提供可理解的安全信息被认为是极为重要的，但本手册未能解决这些问题，因为必须采用专业方法进行评估。将来的进一步发展或许有可能将可理解性测试扩大到适合儿童的方法。

A6.4.1.2 手册中模块 1 和模块 2 的部分介绍了获得代表性样本的建议方法。不应使用大学学生，因为在过去的危险公示研究曾广泛使用过这些学生，而这类人不能被视为本研究中确定的具有代表性的目标群体。

A6.4.2 核心小组

A6.4.2.1 考虑到核心小组的宗旨是确保模块 2 至模块 11 使用的方法适合不同的文化和背景，核心小组的参加者应在被评估的目标群体中尽可能具有代表性。重点应放在来自各种文化的工人和社区成员群体上，而这些文化应不同于制作标签和安全数据单的背景。这将主要适用于农业工人、非农业工人和社区/居民/消费者群体，不论他们有没有文化，这些群体的文化和语言背景有可能使危险信息的传达复杂化。下表 A6.2 就核心小组的类别提出了建议。

A6.4.2.2 建议每个类别至少有两个核心小组。然而，如果某一类别(例如，没有文化的农业工人)的核心小组得到的结果看起来非常类似一个相似小组(例如，没有文化的非农业工人)，那么或许有可能省去下面的小组。只有当测试人相信另外的试验不会出现不同的结果时，才能采取这种做法。一般说来，当不同核心小组得到的结果一致时，建议直接进入主评估(从模块 2 起)。但如果测试结果出现显著差异，或者已经得到的信息不够用于了解其余小组的测试结果，则建议继续组建核心小组，直到获得这样的信息为止。在这种情况下，需要继续进行测试，直到结果一致或清晰为止，可能需要组建比所建议的每个类别 2 个小组更多的小组。

A6.4.2.3 核心小组参加人最好不是参加模块 2 至 11 测试的同一批工人，因为在核心小组内将要组织一些学习。只要可能，各小组都应在语言上具有同源性，因为所有的参加人应该能够至少使用一种共同语言进行交流。

A6.5 问卷和试验设计

A6.5.1 有工作和无工作者的不同子群体经历不同，这会影响他们对危险公示信息的理解。模块 2 至模块 8，以及模块 10 测试不同试验条件下的理解力。对测试人数规模的计算结果，结合后勤支持的

Copyright@United Nations, 2015. All rights reserved

考量，给出了下表 A6.2 所载的最低测试人数。模块 6(标签字体和布局对可理解性的影响)和模块 9(模拟试验)，包括不同标签形式的比较(分别为 8 和 11 层)。因此，对于这些模块来说，需要较多的参加人，以便在每一层内产生足够的案例。其他测试模块(3、4、5、7 和 8)则拥有较少的层(最少 1 层，最多 4 层不等)。因此只需要较少的受测试人。这一方法的使用者可选择将所有模块用于所有的受测试人，在这种情况下，建议的最少受测试人数与表 A6.2 中的模拟 6 和模块 9 建议人数相同。模块 2 和模块 10 则必须按照规定，由所有的参加人完成。

A6.5.2 考虑到全部试验的长度(见表 A6.3)，并出于后勤方面的原因，或许有必要将试验方法进行分解，让不同的受检人只完成某些模块。采取这种方法，就会有更多的人参加试验，但他们只完成评估的某些部分。如果是这样的话，应该记住，所有的受测试人必须完成模块 2 和模块 10, 不论他们已经完成了多少个其他模块。比如，全部模块可细分为几组：

(a) 模块 2、3、8 和 10；

(b) 模块 2、4 和 10；

(c) 模块 2、5、6、7 和 10；

(d) 模块 2 和 11；

(e) 模块 9、2 和 10。

然而，如果可能的话，最好让参加人完成本方法中所包括的所有试验，并对他们付出的努力给予适当的补偿。

表 A6.2：取样规模—建议人数

类别	子类		核心小组：模块 1	测试：模块 2、6 和 10；模拟：模块 9	测试：模块 3、4、5、7 和 8
目标组 1： 工作场所 (a) 管理人员	**群体 1**：生产经理、工程师、技术人员		任选	30-50[a]	25
	群体 2： 工业和农业监督管理人员		任选	30-50[a]	25
(b) 工人	群体： 农业工人	3：有文化	至少一组	100	50
		4：没有文化	至少一组	100[a]	50
	群体： 非农业工人	5：有文化	至少一组	100	50
		6：没有文化	至少一组	100[a]	50
目标组 2：运输	**群体 7**：运输工人		任选	30-50	25
目标组 3： 社区居民/消费者/普通公众	**群体 8**：有文化		至少一组	100	50
	群体 9：没有文化		至少一组	100[a]	50
	群体 10：零售商和分销商		任选	30-50[a]	25
目标组 4：应急人员	**群体 11**：保健专业人员、技术推广人员和应急人员		任选	30-50[a]	25
目标组 5：其他	**群体 12**：执法人员/管理人员		任选	30-50[a]	25

[a] 鉴于组织模拟试验的实际困难，建议在这些组中，只有在资源允许而且切实可行的情况下，才进行模拟试验。

Copyright@United Nations, 2015. All rights reserved

A6.5.3　只要可能，在选择各分小组时，应采用随机取样的方法选择参加人，以尽可能使他们具有代表性。这对于确保结果的普遍性至关重要。即使由于整个试验长度的原因，从同一个分组选择不同的参加人来完成试验的不同部分，参加人的选择也应突出代表性。然而，业已认识到，在实践中，要做到随机选择是十分困难的。尽管如此，仍应牢记，不论采用什么样的选择方法，都应该尽可能使样本具有代表性。

A6.5.4　请注意，在模块内，随机选择小组内受检人至关重要，对此不能将就。随机性对于确保内部比较的有效性也不可或缺，它不同于样本的随机选择，那是确保研究结果具有普遍性所必需的。

A6.5.5　*模拟研究*：由于模拟研究是一种动用资源相对较多的工作，因此建议模拟研究只能在有限的目标群体内进行，可以是农业和非农业工人，也可以是运输人员和消费者。然而，只要资源允许，这样的模拟研究可方便地根据需要用于其他层。

A.6.5.6　感染和共同介入

A6.5.6.1　测试设计要求控制环境。为此，应避免出现这样的情况，即参加人能够看到或听到另一位参加人的试验资料。当对独立变量的掌控是评估的关键环节时，这种情况将使所做的比较失去意义。出现在试验环境中的这种情况被称为感染。

A6.5.6.2　为避免出现感染，在测试进行过程中，参加人应避免相互接触。这可能需要测试小组付出相当大的努力，以确保受检人不会碰到一起。这虽然很困难，但还是应该尽一切努力，尽可能减少感染的可能性。

A6.5.6.3　一个截然不同而又有关联的问题是共同介入，即两个试验组都遇到在试验情景之外发生的介入。共同介入会出现在这样的时候，比如，工厂中的每名工人在测试开始前一周接受详细的危险安全培训。这有可能导致不同危险公示要素的影响受到掩盖，并有可能导致对标签和安全数据单的不同形式的影响估计不足。如果这种情况无法阻止，应注意出现共同介入的可能性。

A6.5.7　分组学习

模块 11 旨在检测在分组学习条件下的可理解程度。它只适用于工人(上表 A6.2 中的群体 3 至 6)，而且需要没有参加模块 2 至模块 8 的工人参试。总共需要对 10 个小组进行测试，其中包括由工厂工人组成的 5 个小组和由农业工人组成的 5 个小组。各个小组应尽可能做到文化水平均衡，有文化和没有文化小组数目大致相等。每个小组均不应超过 10 人，但不少于 6 人。

A6.5.8　关联性

A6.5.8.1　进行可理解性测试时的背景对于含义和理解的准确评估至关重要。对于那些只接受过一点正规教育、全靠望文生义理解危险信息的工人来说更是如此。为此，本方法的大多数测试使用完整的标签，而不是标签或安全数据单的部分内容。接受过良好教育的受检者可能更容易从概念上理解独立的要素，对这些要素的解释可能也与现实世界的学习情景没有多少的联系。正是由于这个原因，所有测试都应该使用真正的标签和安全数据单。

A6.5.8.2　为尽可能提高真实性，应使用容器上粘贴的标签。要将不同的标签贴到各个容器上可能会增加测试者的不必要负担，因此建议将标签贴在标准容器上，测试完后取下。这一程序如果对测试人是过重的负担，可能需要一名助手的协助。重要的是，应将每一种视觉提示提供给受测试人，以最大限度地提高他们的理解可能性，特别是那些没有受过多少正规教育的工人，他们在很大程度上依赖望文生义。因此，不论在任何时间，标签都应粘贴在容器上。使用维可牢尼龙粘带(Velcro)贴在容器上可使这一程序变得简单一些。

A6.5.8.3　为使理解机会标准化，标签上标识的实际化学品将是假的，虽然他们看上去如同真的一样。

Copyright@United Nations, 2015. All rights reserved

这样做的目的是保持关联性，同时又不使不熟悉特定化学品的人处于不利的境地。

A6.5.8.4　如上所述，鼓励用户在试验设计要求的范围内对试验材料进行修改，使材料让当地受检者看上去尽可能可信一些，从而最大限度地提高关联性。

A6.5.9　分项研究的取样数量

分项研究的取样数量是根据双侧α误差 0.1 的 0.8 次方计算得出的，但也考虑后勤的可行性进行了调整。对本方法的初步试验已经证实了这种估计。具体而言，考虑了有选择性地对较少数受检人和目标组进行模拟练习，这在很大程度上是由于预期的后勤限制。

A6.5.10　翻译

A6.5.10.1　语言是许多危险公示的关键。虽然本方法尽可能考虑到语言差异，但如果翻译不当和不标准，可能使试验出现很大的误差。因此，应该特别注意翻译的准确性。应遵循的程序如下：

(a)　让两个英语流利(本方法使用的语言)的人独立地将问卷翻译成标引语言(目标组使用的语言)；

(b)　然后再由另外两名翻译各自独立地将这两个译文翻译回英文，而且这两个人不是最初的两名翻译。

A6.5.10.2　反向翻译的目的是实现第一轮的翻译错误少于 5%。对译文中的错误进行澄清，应纠正含糊不清之处。在可能的情况下，应设法将两篇译文中所有正确翻译的内容合并起来并作反向翻译。

A6.5.10.3　如果无法做到后者，可将错误率较低的译文作为首选译文。如果错误率超过 5%，则需要进行第二轮反向翻译。

A6.5.11　测试和核心小组的时间安排

A6.5.11.1　进行测试和组建核心小组，必须选在对被测试者及其雇主(如果适用的话)都方便的时间。在关键的农忙季节(比如播种、耕地、喷洒，或收割)，不应要求农业工人参加测试。测试工人应在工作时间进行，而且不应让他们因参加测试而遭受经济损失。不建议在没有适当补偿的情况下，让工人利用他们自己的时间(午饭时间或下班以后)参加测试。如果工人同意在午餐期间参加测试，时间长短必须合适，并给予适当的补偿(返回时间，提供午餐等)。

A6.5.11.2　表 A6.3 根据在两个南非工厂中进行的初步试验的结果，给出了完成每个模块所需要的估计时间。取决于模块的不同和模块使用人员的熟练程度，整个测试时间在 20 分钟到 2 个小时不等。如果工人没有什么文化，测试时间需要长一些。

表 A6.3：危险公示可理解性测试的大致时间

模块	时间(分钟)
1	60 - 120
2	30 - 45
3	45 - 75
4	75 - 105
5	20 - 30
6	20 - 30
7	20 - 30
8	45 - 75
9	30
10	30 - 45
11	120 - 180

Copyright@United Nations, 2015. All rights reserved

A6.5.12 回答评分和编码

A6.5.12.1 可理解性测试的回答评分要求专家对回答的正确性作出判断。过去在津巴布韦的经验表明，当测试人尽可能使他们的方法标准化时，自由回答的内容分析可能是可行的。

A6.5.12.2 本方法要求有一组专家在场，以便对理解程度进行所需要的评分。专家小组应在开始研究之前按下列步骤确定：

(a) 选择一组有各种经验的专家，包括(一名或多名)雇员、雇主和从业者，以及在编码和评分方面技能熟练的研究人员；

(b) 召集一次由专家小组参加的讲习班，对每个模块中所列的问题，审查可能做出的回答的性质。审查全球统一制度程序的文件，以便就什么样的回答可归为以下类别达成一致意见：

(一) 正确：意思与全球统一制度结构的构想相同或者完全一致。这包括虽不与全球统一制度的意思 100%相同，但足以作为安全行动或防护措施基础的回答；

(二) 部分正确：某些意思是正确的，但不足以确保适当的安全行动或防护；

(三) 不正确：给出的意思完全错误，或者与全球统一制度要表达的意思风马牛不相及；

(四) 意思相反(严重混淆)：给出的意思不仅不正确，而且理解与全球统一制度的意图恰好相反。这种严重混淆有可能导致危险的行为或行动；

(五) 不能回答/不知道；

(c) 在 5 或 10 名受测试人中试行本问卷。对照所选择的标准检查结果；

(d) 如果结果表明存在显著差异，重复上述过程，直到就标准达成一致为止。

A6.5.12.3 对不同模块中的问题，视情况在每一模块下讨论对回答作进一步编码。

A6.5.13 分析

建议对这些模块进行的分析，是对照不同层进行的比例和办法的简单计算。可进行更为复杂的分析，这些分析见各个模块。可尝试对可理解性进行总体估计，办法是将不同层的受测试人的结果综合在一起，但也应根据不同层的权重和已知影响可理解性的其他人口统计学因素的权重进行调整。

A6.5.14 反馈和后续步骤

应让所有的受测试人有机会看到可理解性评估结果，并对面试和测试程序作出反馈。

A6.5.15 后续评估

应在一个月和一年以后对参加这些评估的受测试人再次进行面试，以考察他们对当时接触到的全球统一制度危险信息还能记住多少，以及所产生的中长期影响。根据资源和后勤情况，可不对在基线完成的所有模块进行重新测试。

Copyright@United Nations, 2015. All rights reserved