外围声化中的声音生态感知研究

Maxime Poret
波尔多大学、波尔多国立高等电子、计算机、无线电通信与机械工程学院、法国国家科学研究中心、LaBRI、UMR5800，法国塔朗斯 F-33400 maxime.poret@labri.fr

Catherine Semal
波尔多大学、波尔多国立高等电子、计算机、无线电通信与机械工程学院、法国国家科学研究中心、神经科学综合研究所、UMR5287，法国波尔多 F-33076 catherine.semal@ensc.fr

Myriam Desainte-Catherine波尔多大学、波尔多国立高等电子、计算机、无线电通信与机械工程学院、法国国家科学研究中心、LaBRI、UMR5800，法国塔朗斯 F-33400 myriam@labri.fr

arXiv:2302.03957v1 [cs.HC] 8 Feb 2023

摘要

基于一项关于 3D打印 的案例研究，我们一直在试验将多维数据声音化，用于外围过程监控。在先前的一篇论文中，我们测试了一种声景的有效性，该声景将刻意不协调的自然声音与音乐声音结合起来。这一做法基于这样一个假设：听觉刺激之间的生态连贯性越低，就越容易彼此凸显，从而提高对各种通知的反应率。在本文中，我们通过测试两种新的声学生态来进一步验证这一假设，它们分别只由音乐声音或自然声音构成。随后，我们让这两种声学生态经历与先前研究相同的双任务评估流程，以便进行比较。结果似乎支持我们的假设，因为对新的声学生态的检测准确度不如原先的声景。不过，测试者似乎认为自然声音集合的干扰性较低，也有助于提升外部主任务的表现。我们希望这项研究能成为更大规模研究文献的一部分，最终为声景设计中生态连贯性的影响提供更明确的答案。

1. 引言

通过非语言声音呈现数据，即声音化 [1], ，已被证明能够为过程监控任务提供多种适当的通知。Gaver等人 [2], 随后Rauterberg等人 [3], 的研究表明，即便监控指标数量庞大，这种方式也同样适用。尽管数十年来，人们一直使用它在工业流程中传达警报（通常采用铃声、哨声、鸣笛声和蜂鸣声等令人过度紧张的形式），但近年来更先进的声音与数据处理技术的出现，促使人们努力摆脱“宁可过度谨慎也不要冒险”的范式 [4]，转向更平和、侵扰性更低的声景 [5]。本文聚焦于“外围式”声音化：它应让听众能够获取有关数据状态的各种听觉通知，同时也允许他们处理更需要注意力的主要任务 [6]。

版权：© *2022*年 *Maxime Poret* 等人。本文为开放获取文章，依据知识共享署名 *3.0* 未移植版许可协议发布，该协议允许任何人以任何媒介不受限制地使用、分发和复制本文，但须注明原作者和来源。

传统上，这类声音化的设计指南建议使用少量同时播放、持续发声的声音流 [7], ，并在选择时考虑生态连贯性。这意味着既要确保这些声音流不会被使用情境中原本存在的声音掩蔽，也要确保为声音分配的频谱带宽能够避免它们彼此掩蔽 [3, 6, 8]。我们的目的并非质疑这些防止掩蔽的预防措施，因为它们似乎不言自明。然而，声音化设计者往往似乎还遵循一条额外规则：这种连贯性也应特别注重声音选择在逻辑上的合理性，因此声景通常围绕某些主题（或“声学生态”）展开，例如“森林”“海滩”，或各种纯音乐形式的呈现。

近期关于听觉处理的研究表明，声音的辨别发生在神经生理层面，并通过一系列并行分类过程实现 [9–11],；即使自然声音经过大幅缩短和简化，仍能在不协调的音乐声音集合中被识别出来 [12]。因此，我们不禁想知道，选择属于截然不同认知类别的声音（例如音乐声音与动物叫声）是否更有利，从而使皮层层面的分类过程能够更好地区分它们。

在此前一项关于过程监控声音化的研究中 [13],，我们提出假设：与逻辑上连贯的刺激集合相比，不协调的刺激集合或许能让人更快、更准确地辨别通知。本研究旨在开展一项实验来检验这一假设。

对于研究过程监测声音化的学者而言，基于不同设计策略提出多组刺激是相当常见的做法（通常是在听觉图标与音乐映射之间进行二分） [14–18]。然而，包含音乐声音与自然声音这两者的“混合”生态的研究似乎较为少见 [8, 19]；据我们所知，也没有研究对这些设计选择进行超越定性用户调查的完整比较评估。

在本研究中，我们设计并评估了两组符合生态特征的刺激，以便将它们与此前测试过的一组不符合生态特征的刺激进行比较。为此，我们沿用了部分现有声音流，并补充了新的声音流；这些新声音流是根据新组各自的生态指南（第 3 节）选取的。

随后，我们采用与此前相同的双任务实验（第 4 节）对它们进行评估，以便比较每种生态下的异常标注质量、对主要任务表现的影响，以及听众的审美评价（第 5 节）。

2. 背景

2.1 3D打印

本项目源于一项案例研究，探讨如何利用声音化来帮助监控电弧丝材增材制造过程 [20]。在这一过程中，通过将材料逐层添加来构建金属零件；材料在高压电流通过导线产生的电弧作用下受热、熔化并沉积。由于这一过程往往噪声大、闪光强，而且总体而言令人不适，不适合在旁边停留（更不用说直接监控），因此，过去几年一直有人尝试利用其他模态来增强对其的观察，例如触觉振动 [21], 、视觉增强现实 [22] ，或者在我们的案例中使用外围听觉显示 [13, 23]。

预计该显示的外围特性能够为操作员提供一种更远程的监控方式，使他们可以离开打印机，处理其他任务，例如查看电子邮件、管理日程、操作其他机器等，同时仍能在收到通知后返回并对该过程作出响应。

下文介绍了我们的显示器应帮助监测的制造标准。阈值和公差的完整详情另请参见表1的第一列。

- 熔池尺寸（WPD）是打印头正下方材料沉积物的宽度和高度。
- 零件高度（PH）是打印头下方所有已构建层的累积高度。
- 熔池温度（WPT）是材料沉积物的局部温度。
- 零件温度（PT）是正在构建的零件的平均温度。

2.2 模拟过程

截至本文撰写时，能够提供这些监控数据的传感器仍在研究中[24]。因此，我们的声音化是通过在合成日志文件上运行 Su- perCollider [1] 脚本生成的，这些文件模拟了打印过程中各项标准应有的表现。

为了恰当地评估声音与标准之间的关联，需要考虑用户在适当培训阶段中的学习曲线，因此实验需要持续较长时间（每周进行数次试验，持续一两个月）。理想情况下，这还需要招募一批打印机操作员，他们已经熟悉打印流程及其相关条件和标准。

[1] https://supercollider.github.io/

就本项目目前的进展而言，这种实地实验尚不可行。因此，我们决定暂不研究声音隐喻与制造标准之间的心理关联，而是将实验重点放在声音本身的检测，以及声音在（模拟的）外围工作环境中的可用性上。尽管存在这一局限，我们仍希望为未来的大规模实验奠定基础，以便研究在真实监控情境中使用类似显示界面的情况。

3. 声音映射

3.1 一般原则

尽管本文并未研究声音与各项标准之间的关联，但我们希望通过仔细选择与不同类型物理数据相关的隐喻，为下一步工作做好准备。

对于自然刺激，我们避免选择那些在没有任何视觉背景的情况下难以区分的同时声音。例如，大雨、沸水或流动的河流的声音很容易彼此混淆，或者在同时出现时相互掩盖。同样，在选择音乐刺激时，我们确保各个声流具有不重叠的音高范围，以及不同的音色和粒度。

音色、乐器或声源的选择因生态而异，但我们力求在所有生态中尽可能保留多项心理声学考量。主题维度是此处的主要变量。因此，声景始终包含一个“空闲”状态，由两种持续刺激构成，分别指示正常的 WPD 和 PH 值。WPD 的两个维度都映射到三个参数（高度对应音高，宽度对应节奏，而两者都对应响度），共同构成缓慢重复的声音爆发，隐喻性地表现物质的逐渐累积。PH 通过持续的声音传达，该声音可演变为三种状态（正常、过低、过高），代表物质沉积其上的基底。正常情况下，WPT 以静默传达，但可根据异常的方向以两种形式之一出现。事实上，没有必要不断提醒听众温度正常，尤其是当我们希望在没有异常时避免显示过载。PT 通过单个声音事件传达，用以表示阈值已被越过。这是一种最后手段警报，我们选择让它在概念上与 WPT 的声音相关联，因为预计全局温度警报会由持续较长时间的局部温度异常触发。

以下小节将分别介绍每种生态的映射选择。表1概述了每种刺激的映射方式，以及这些映射如何组合形成三种生态。

3.2 混合生态

如前文所述， [13], 混合生态的设计旨在通过两种不同类别的声音，分别传达几何（WPD、PH）和热学（WPT、PT）标准

WPD 宽度： $4mm \pm 10\%$ 高度： $3mm \pm 10\%$	高度差	主音音高：C5（正常）至 F6（最差）	琶音（混合与合成器）
	两种差异	响度：0.02x 至 0.2x	
	宽度差	偏移间隔：0.5 至 1.5 s	
	高度差	播放速度：0.5x 至 2.0x	水滴（自然）
	两种差异	响度：0.1x 至 0.8x	
	宽度差	间隔：0.5 至 1.5 s	
PH 当前图层 $\pm 1.5mm$	相对差异	音高：A3 ±3 音	持续音（混合与合成器）
	绝对差异	响度：0.1x 到 0.4x	
	相对差异	混音：鸭子、其他或乌鸦	鸟鸣（自然）
	绝对差异	响度：0.0x 至 0.3x	
WPT $2000°C \pm 10\%$	差异极性	音高：220 Hz 或 880 Hz	叮当声（合成器）
	绝对差异	响度：0.0x , 0.1x 或 0.2x	
	差异极性	选择：噼啪声或沸腾声	水（混合与自然）
	绝对差异	响度：0.0x , 0.2x 或 0.5x	
PT $600°C$	阈值	触发：随机成分	铃声（合成器）
		触发：声音样本	滋滋声（混合与自然）

表1. 各监测标准（左）与刺激（右）之间映射关系的概述。每个刺激名称后括号内为其所属声学生态的名称。

这些声音分别基于听觉图标 [25]和抽象耳标 [26] 指南。

热学标准相对容易通过隐喻性听觉图标来传达，因为现实世界中有许多现象，其声音与温度的概念直接相关（沸水、冰块噼啪声、燃烧的火、滋滋作响的水）。然而，几何标准则需要更高层次的抽象，才能通过声音传达。事实上，很难将自然现象与“宽度”或“高度”的概念直接联系起来。因此，我们选择将这些维度映射到音乐流中的感知参数。

因此，混合生态由以下4种刺激组成：琶音（WPD）、持续音（PH）、水（WPT）和滋滋声（PT），详见下文。

琶音是一段由大调音阶中的3个音高（主音、三音和五音）构成的循环乐句，由默认的 SuperCollider 合成器演奏，音色简单，类似钢琴。在空闲状态下，这段声音流由低音量的 C5大调琶音重复构成，每个音符之间间隔1.5秒。当 WPD 的波动改变主音音高、时长或响度时（映射关系见表1），新定义的乐句会打断先前开始演奏的乐句。

持续音是一种连续播放的带通滤波锯齿波，其音高不断波动。在空闲状态下，它会以低音量持续播放一个 A3音符。此后，其音高可根据 PH 偏差的方向降低或升高。

水可以看作一小片水体，大部分时间处于静止、无声的状态，但温度发生异常快速变化时，它可能开始沸腾或冻结。

作为这一水隐喻的延续，滋滋声警报由一种更猛烈的蒸发声构成，让人联想到水被泼到烧得通红的表面上。

3.3 合成器生态

对于合成器生态，我们只选择了音乐声音，采用多种合成音色，并将其抽象地映射到感知参数上。我们沿用为混合定义的琶音和持续音，并为热维度 WPT 和 PT 设计了两种新的刺激。

对于 WPT，我们不再选择自然隐喻，而是将异常映射到快速打击乐声粒的音高上（以 0.06 秒正弦波脉冲重复播放，并采用快速起音和衰减），并将其称为“叮当声”。该刺激在空闲状态下静音。异常出现时，其响度随之增强，音高则有两种可能：过热时为 220 Hz，过热时为 880 Hz。

PT阈值警报被“铃声”取代，这是一种非谐和声音，起音迅速、衰减缓慢，让人联想到敲击大型铃铛的声音。与先前的滋滋声和水一样，这一隐喻旨在延续叮当声的音色。为了增强听众对该警报的新奇感，我们使其音色部分由随机生成：在恒定的 440 Hz 基频周围，随机选取 3 个频率介于 220 至 880 Hz 之间的非谐波分量。

3.4 自然生态主题

自然生态主题基于听觉图标设计原则。我们沿用先前为混合主题定义的水和滋滋声听觉图标，并为几何维度 WPD 和 PH 新增了两种刺激，采用类似于“森林池塘”的生态主题隐喻。

因此，自然生态主题由水滴（WPD）、鸟鸣（PH）、水（WPT）和滋滋声（PT）这些刺激组成。

正如在“混合生态”的描述中所指出的，很难选择能够与某个物体的几何维度相关联的自然声音隐喻。不过，我们可以利用自然声音处理的一项特性，

Gaver 将其称为“日常聆听” [27], , 即听者可以根据声音的感知属性，辨识出产生该声音的物体或动作的特征。例如，较大的物体落入池塘时，会产生比小物体更低沉、持续时间更长的“溅水”声。我们正是依据类似的逻辑定义了水滴刺激：使用水滴落到表面上的短声音样本来传达 WPD。高度被映射到其播放速率，因此当焊池变得更高时，声音听起来会具有更高的音高（仿佛下落得更猛烈）；宽度则被映射到其出现速率，因此当焊池变得更宽时，声音会被听到得更加频繁（仿佛发生了溢流）。

传达 PH 的鸟鸣刺激基于这样一种构想：构建一层持续不断的杂鸟鸣声，使得在闲置状态下，整体悦耳而熟悉的噪声中不会有任何信息特别突出。出现异常时，某一物种的声音便会显现，从而打破噪声并引起听者的注意。该物种可能是鸭子（PH 过高时），也可能是乌鸦（PH 过低时）。

这种以鸭子或乌鸦来传达高度的抽象隐喻选择，旨在将这一维度映射到西方民间传说中这两种鸟各自的情感内涵：鸭子是人们熟悉的家禽（评价正面，因此具有“高”极性），而乌鸦是食腐动物，也是厄兆（评价负面，因此具有“低”极性）。

4. 实验

为保持实验条件的可比性，同时考虑到该游戏能对周边监测的工作情境进行基础模拟，我们沿用了最初为评估混合生态而开发的同一款双任务游戏。实验界面的截图见图1。主要任务是使用鼠标完成一项非常简单的序列复制任务；次要任务则是在声景中聆听通知，并在通知出现时勾选侧边的复选框进行标注。

由于实验期间的卫生限制，评估游戏必须以交互式网页的形式发布，并使用 JavaScript 和 PHP 元素记录测试期间的用户操作。每位参与者都在自己的环境和设备上完成测试。我们承认这可能导致测试条件不统一，并希望通过足够大的测试者样本来弥补这一点。实验的两种版本各采用一种新的声学生态，均实现在同一页面上；用户连接后，系统会根据目前已完成测试中哪种刺激出现次数较少，选择相应的一组刺激。如果两种声学生态的测试次数相同，则随机选择一种。

参与者通过训练阶段开始实验，在此期间，他们会了解主要任务，随后分别听取每种声音刺激，最后聆听一些处于空闲状态和异常状态下的完整声景示例。最后，他们必须成功完成两个关卡，以确保理解了整个流程，然后才能进入正式实验。正式实验由十个各30秒的关卡组成

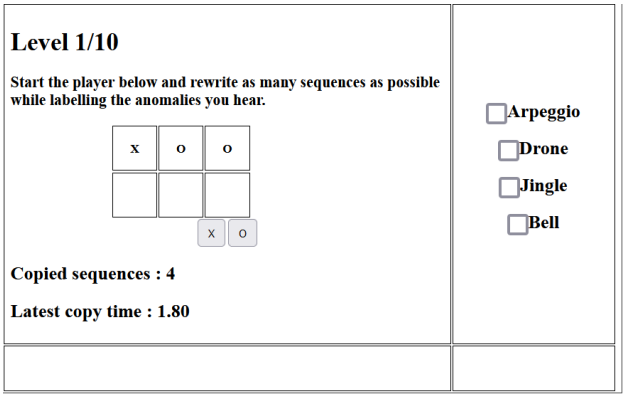

图1. 实验界面在某一关卡中的屏幕截图，此处展示的是合成器生态。画面中央，玩家通过依次点击“X”和“O”按钮，重现上方一行显示的序列。每完成一个序列，就会出现一个新序列。右侧的方框可供玩家标记出现的异常。这些方框上的标签取决于所测试生态中所用刺激的名称。

其中包含以随机顺序选取的各种异常组合。

应用会记录在关卡过程中勾选异常框的时间，以及每个序列的大小和完成时间。

完成实验后，参与者会被引导至一份简短问卷。问卷首先询问他们通常的人口统计信息（年龄和性别），随后请他们针对一系列与练习有关的陈述（例如“任务很难”或“声音令人紧张”，见表2）作答，选项为“不同意”“有些同意”和“同意”。最后，自由评论框允许他们进一步补充反馈。

5. 结果

合成器由15名参与者进行测试。其中4人曾参加过早期版本的实验。8名男性、5名女性、2人未注明。年龄为20至69岁（$\mu = 35.9$, $\sigma = 15.9$）。

自然由16名参与者进行测试。其中2人曾参加过先前版本的实验。7名男性、9名女性。年龄为19至62岁（$\mu = 24.75, \sigma = 10.5$）。

在前一项实验中，混合共由43名参与者进行测试，其中20名男性、23名女性。年龄为18至67岁（$\mu = 32.3$, $\sigma = 12.4$）。

我们使用实验期间收集的数据，为各显示界面计算了一系列质量指标，详见以下小节：每种刺激的标注准确率、标记异常所需的反应时间、主要任务表现的分布，以及定性用户反馈。

5.1 标注准确率

信号检测理论 [28] 使我们能够利用标注任务中每位参与者的命中率（H）和虚报率（FA），计算一个名为敏感度指数的指标，即 d',

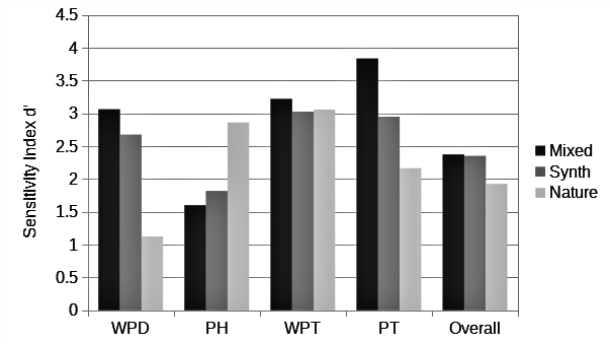

图2. 各刺激类型（横向分区）和各生态（颜色）的平均敏感度指数。

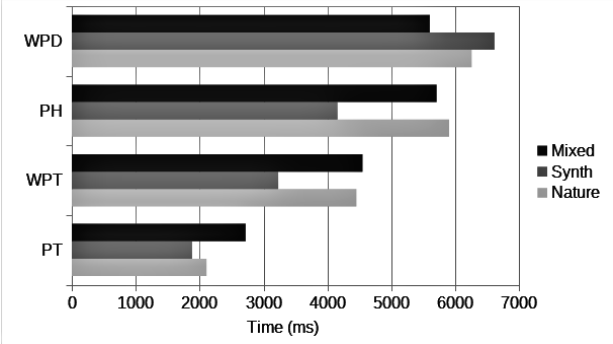

图3. 各刺激类型（纵向分区）及各声学生态（颜色）对应的平均标注时间（毫秒）。

因此 $d' = Z(H) - Z(FA)$，其中 Z 为累积正态分布的反函数。我们采用“平均敏感度”策略，对每种刺激类型下所有参与者的 d' 指数取平均。由于部分用户的比率过于“精确”，无法进行 Z 计算（总比率为 0 或 1），我们将其近似调整为 10^{-2} 实际值（分别为 0.01 和 0.99）。这相当于认为，即使每种刺激进行一百次试验，效率最高的测试者仍有 1% 的概率出错。

该敏感度指数的数值越高，表示相应刺激的可辨识度越强。由于它没有维度或刻度，因此无法绝对衡量信号的质量（只能说明带有 $d' > 1$ 的信号被正确识别的概率高于随机水平），但可用于比较在同一实验中评估的不同信号。

只有在异常开始后点击标注复选框，且一直保持勾选状态直至关卡结束，我们才将该标注视为命中。每次对异常的“预测”都视为假阳性，每次“改变想法”则计为未标注。

请参见图2，其中按生态分别展示了各刺激及总体的平均敏感度指数直方图。

- WPD：水滴的效率明显低于琶音。琶音在混合中的表现略优于合成器。
- PH：鸟鸣明显比持续音更高效。
- WPT：三种声学生态之间没有明显差异，不过水在混合声学生态中的表现似乎略优于自然声学生态。
- PT：滋滋声在混合声学生态中的检测效率远高于自然声学生态。
- 总体而言：混合声学生态略优于合成器声学生态，而自然声学生态的效率明显低于两者。

5.2 标注时间

在图3中，我们绘制了每种刺激类型和每种生态下的平均标注时间。同样，这些数据仅根据参与者在异常出现后正确识别异常的情况计算。比较这些数值后，我们注意到，合成器似乎总体上使反应时间更短，但 WPD 除外：在 WPD 中，作为混合的一部分，琶音刺激的标注速度比作为合成器的一部分更快。

5.3 主要任务表现

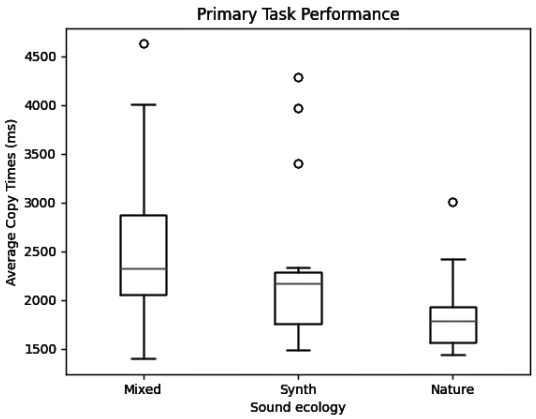

图 4. 各种生态条件下测试者平均序列复制时间（毫秒）的箱线图。

我们计算了每位参与者复制一个序列所需的平均时间，并将其作为衡量他们在主要任务中表现的指标。将各生态条件下这些分数的分布绘制为箱线图（见图 4），便能呈现测试者在每种条件下的整体表现。在这些图中，数值越低表示成绩越好，因为要达到最佳表现，复制时间应尽可能短。

总体而言，自然条件下参与者的表现似乎最好，其次是合成器条件。混合条件下的分数分布更广，因为该实验的参与者人数要多得多。这似乎表明，混合条件下的主要任务表现略逊于另外两种条件。

5.4 定性调查

评估结束时的问卷由七项陈述组成（其中三项涉及声音设计，四项涉及与声音相关的主要任务），

	混合	合成器	自然
声音容易区分	71.43	82.14	**83.33**
声音会分散对任务的注意力	7.14	**25.00**	14.29
声音令人感到压力	19.05	**42.86**	30.00
任务令人感到压力	16.67	**25.00**	16.67
任务很有趣	66.67	46.43	**73.33**
任务很困难	14.29	**21.43**	13.33
任务会分散对声音的注意力	11.90	**21.43**	3.33

表2. 定性调查回答概览。每条陈述中，同意率最高的部分以粗体标出。

参与者可以使用三档量表作答，从“不同意”到“同意”。为分析回答，我们为该同意程度量表的每一档赋予相应数值：“不同意”计为 0，“比较同意”计为 50，“同意”计为 100。随后，我们对每项陈述的数值取平均，得到一组百分比，用以表示总体同意率。回答概览见表 2。

总体而言，合成器似乎比另外两种声学生态更令人感到压力（第2–7行）。这甚至包括测试者在听到这组声音时，认为主要任务更令人感到压力的情况（第4行）。

在一条关于综合生态声景的自由评论中，一位测试者表示，持续音刺激过于敏感。这个问题在最初的混合评估中也曾被报告。这似乎表明，持续音的大多数错误可能是由误报而非漏报造成的。

测试自然声景的参与者似乎认为其刺激集更容易区分（第 1 行），尽管他们在标注任务中的表现较差。这可能是因为，虽然声音类型本身很容易区分，但出现异常时的调制可能更难察觉。两位参与者报告说，他们觉得很难区分水滴的正常与异常情况；这一点显然反映在其表现中，而且很可能不止报告此问题的两位参与者遇到了这种情况。

对于自然，四名测试者表示难以区分滋滋声和水声；此前也有报告指出，混合存在同样的问题。从某种程度上说，这种混淆令人意外，因为选择滋滋声时，特意将其设定在比水声更高的频率范围，并赋予其极具特征性的瞬态包络，而不是连续的冒泡声。尽管单从感知层面看，这两种声音本应截然不同，但日常聆听似乎在周边监测中占了上风，两者都被简单地听成了“沸水的声音”。此外，由于滋滋声是一种最后手段警报，在实验中出现的频率低于其他刺激，这可能导致测试者在训练期间初次接触之后便将其遗忘。我们猜测，适当的长期训练或许能大幅减少区分水声和滋滋声时遇到的问题。此外，我们注意到，这两种刺激的敏感度指标仍优于混合中的任何其他刺激。

混合声景中使用的声音相对更常被评为难以区分（第 1 行）。这可能是前述滋滋声与水的混淆，以及持续音过度敏感共同作用的结果。

我们注意到，包含自然声音的声学生态往往获得更积极的评价：

- 自然声音刺激被认为更容易区分（第 1 行）。
- 混合最不容易分散注意力，也最不令人感到紧张（第 2、3 行）。
- 在对主要任务的干扰程度上，混合与自然并列为最不令人感到紧张的选项（第 4 行）。
- 使用自然声音时，主要任务被认为更有趣、难度也更低（第 5、6 行）。
- 在处理主要任务的同时，自然声音更容易跟随（第 7 行）。

最后这几点似乎证实了在监测显示界面中使用听觉图标的常见论据。

最后，参与自然主题的三位参与者表示，他们觉得用于水主题的冰块噼啪声令人分心、苦恼或有压力。此前在混合主题中未曾报告过这一问题，我们怀疑这可能是自然主题中所用声音的主题语境带来的一种新型情感副作用。

6. 讨论

尽管我们的结果似乎表明了一些与假设相关的趋势，但应注意，统计分析（两两单因素方差分析）显示，这些趋势并不具有充分的代表性，因此解读时应保持谨慎。

自然生态主题的标注准确率最低，但测试者似乎对其主要任务的评价更积极，且在该任务上的表现也更好。这可能与测试者人口统计特征造成的偏差有关：测试自然主题的参与者总体上比测试合成器主题的参与者年轻（其中15人年龄在40岁以下，而合成器主题组为8人），因此他们可能普遍更擅长基于鼠标的任务。然而，这种解释无法说明为何合成器主题的反应时间总体上短于自然主题。或许合成器主题和自然主题的测试者只是分别优先关注双重任务的不同方面，即刺激的标注或序列复制。尽管如此，定性反馈似乎印证了人们常提到的听众对自然音效的偏好，而这种偏好或许反映在他们的主要任务表现上，使监控体验更加轻松。

测试者对合成器的声音辨识度更高，却更不喜欢它；这可能是因为它更接近传统的“功能性”警报模式。这种模式能够发出强烈的通知，但长期使用往往会带来更大的压力。

我们注意到，在不同声学生态中共享的一些刺激，在混合生态中呈现时比在各自连贯的生态中更容易被辨识：琶音在混合生态中比在合成器生态中更有效，而滋滋声和水声在混合生态中也比在自然生态中更有效。这似乎证实了我们最初的假设：设计一种生态，让不同类别的声音能够更好地区分开来。

7. 结论

作为探索在3D打印过程监控中使用外围声化的一部分，我们研究了声景中生态连贯性的影响。我们进行的比较评估似乎支持了我们的假设：不连贯的生态有助于更好地检测和辨别刺激。不过，我们也注意到，用户认为听起来自然的生态侵扰性较低，而听起来合成的生态则带来了更短的反应时间。就这些标准而言，混合生态处于另外两者之间，似乎兼顾了低侵扰性和快速反应。

我们注意到，尽管自然生态主题在刺激辨别方面的整体表现不如其他主题，但作为 PH 通知的刺激，鸟鸣的表现远胜于持续音。假设不连贯的生态设计值得进一步研究，我们现在想知道，若加入这种自然声音，是否能构建出更好的混合生态。更一般地说，为了尽可能全面地确定哪种刺激组合能带来最佳反应率，我们或许需要对所有组合进行实验。这也能让我们检验，如果 WPT 和 PT 分别由完全不同类型的声音传达（水 + 铃声，或叮当声 +滋滋声），是否能更准确地检测到它们。

在这项工作中，我们设计了若干组刺激，认为它们能够代表对音乐声音或自然声音进行连贯的生态考量。尽管我们开展的实验似乎更倾向于不连贯的生态，但我们始终谨慎地认识到，这些结果很可能反映的是我们自身特定的设计选择，而非声音生态环境本身的类型。因此，我们希望未来能出现更多比较连贯与不连贯声景设计的实验，并据此对通知系统中的生态不协调问题得出更明确的结论。

致谢

本文作者谨此感谢波尔多大学的 SCRIME 为本研究提供场地，同时感谢 Addimadour平台和 Ana¨ıs Domergue 允许我们亲自观察电弧丝材增材制造过程，并回答我们关于该主题的问题。我们还要感谢 Nadine Couture 和 S´ebastien Ibarboure（ESTIA）最初向我们提交这一案例研究，并积极参与本研究前几版的校对和文档编写工作。最后，我们要感谢 Matthias Robine（LaBRI) 以及 Emmanuel Duc（SIGMA Clermont）为声音设计和评估流程的选择提供的见解。

8. 参考文献

[1] G. Kramer、B. N. Walker、T. Bonebright、P. Cook、J. H. Flowers 和 N. Miner，“声音化报告：领域现状与研究议程。为美国国家科学基金会撰写的报告，由国际听觉显示社区成员编写”，载于国际听觉显示社区 *(ICAD)*，新墨西哥州圣菲，1999 年。[2] W. W. Gaver、R. B. Smith 和 T. O’Shea，“复杂系统中的有效声音：ARKola 仿真”，载于计算机系统中的人为因素 *SIGCHI* 会议论文集，1991 年，第 85–90 页。[3]M. Rauterberg 和 E. Styger，“声音反馈在操作工厂模拟器过程中的积极作用”，载于国际人机交互会议。施普林格，1994 年，第 35–44 页。[4] R. D. Patterson、T. F. Mayfield、D. E. Broadbent、A. D. Baddeley 和 J. Reason，“工作环境中的听觉警告音”，《伦敦皇家学会哲学汇刊》*B*辑：生物科学，第 327 卷，第 1241 期，第 485–492 页，1990 年。[5] M. Weiser 和 J. S. Brown，“平静技术的时代即将到来”，1996 年。[6] P. Vickers，“用于过程监控的声音化”，载于《声音化手册》，T. Her mann、A. Hunt 和 J. Nuehoff 编。Logos 出版社，2011 年，第 18 章，第 455–491 页。[7]T. Hildebrandt、T. Hermann 和 S. Rinderle-Ma，“连续声音化提升外围过程监控中交互的适切性”，《国际人机交互研究杂志》，第 95 卷，第 54–65 页，2016 年。 [在线]。可访问：
h
t
t
p
s://
w
w
w.
s
c
i
e
ncedirect.co
m/science/article/pii/S107158191630074X[8] E. D. Mynatt、M. Back、R. Want、M. Baer 和 J. B. Ellis，“音频光环的设计”，载于计算机系统中的人为因素 *SIGCHI* 会议论文集，1998 年，第 566–573 页。[9] M. M. Murray、C. Camen、S. L. G. Andino、P. Bovet 和 S. Clarke，“大脑快速辨别物体声音”，《神经科学杂志》，第 26 卷，第 4 期，第 1293–1302 页，2006 年。

[10] S. G. Lomber 和 S. Malhotra，“听觉皮层中‘什么’与‘哪里’加工的双重分离”，《自然-神经科学》，第 11 卷，第 5 期，第 609–616 页，2008年。

[11] A. M. Leaver 和 J. P. Rauschecker，《自然复杂声音的皮层表征：声学特征和听觉对象类别的影响》，

V. Isnard、M. Taffou、I. Viaud-Delmon 和 C. Suied，《听觉草图：极其稀疏的声音表征仍可被识别》，*PLOSONE*，第 11 卷，第 3 期，第 1–15 页，2016 年 3 月。 [在线]。可访问：https://doi.org/10.1371/journal.pone.0150313[13] M. Poret、S. Ibarboure、M. Desainte-Catherine、C. Se- mal 和 N. Couture，《用于3D打印过程监测的外围听觉显示》，载于*2021*年声音与音乐计算会议论文集（*SMC 2021*），意大利都灵，2021年。[14] J. Cohen，《“柯克在此”：使用类型声音监控后台活动》，载于计算机系统中的人为因素*INTERACT ’93*和*CHI ’93*会议伴随论文集，系列：CHI ’93。美国纽约州纽约：计算机协会，1993年，第63–64页。 [在线]。可访问：https://doi.org/10.1145/ 259964.260073[15] M. Gilfix 和 A. L. Couch，《Peep（网络声音监视器）：用声音监控网络》，载于第*14*届*USENIX*系统管理会议论文集（*LISA’00*）。美国加利福尼亚州伯克利：USENIX协会，2000年，第109–118页。[16] T. Hermann、T. Hildebrandt、P. Langeslag 和 S. Rinderle-Ma，《优化用于外围过程监控的声音化之美学与精确度》，载于第*21*届国际听觉显示会议论文集（*ICAD 2015*），奥地利格拉茨，2015年，第317–318页。[17] S. Lenzi、T. Riccardo、T. Ginevra、S. Galelli、P. Ciuc- carelli 等，《揭示水分配系统中的网络攻击：威胁和异常数据声音化的实验方法》，载于第*25*届国际听觉显示会议。国际听觉显示社区，2019年，第125–132页。[18] A. L. Aldana Blanco、S. Grautoff 和 T. Hermann，《支持心肌梗死诊断与监控的心电图声音化》，多模态用户界面期刊，第14卷，第2期，第207–218页，2020年。[19] S. Matinfar、T. Hermann、M. Seibold、P. F¨urnstahl、M. Farshad 和 N. Navab，《高敏感度外科任务中的过程监控声音化》，载于*2019*年北欧声音与音乐计算会议论文集（*Nordic SMC 2019*），2019年。[20] S. W. Williams、F. Martina、A. C. Addison、J. Ding、G. Pardal 和 P. Colegrove，《电弧增材制造》，材料科学与技术，第32卷，第7期，第641–647页，2016年。 [在线]。可访问：https://doi.org/10.1179/1743284715Y.0000000073[21] S. Ibarboure，《通过振动触觉反馈感知信息，以更灵活地实施机器人增材制造工艺》。博士论文，波尔多大学，2021年6月。[在线]。获取地址：https://tel.archives-ouvertes.fr/ tel-03361862[22] A. Ceruti、A. Liverani 和 T. Bombardi，《3D打印过程中的增强视觉与交互式监控》，发表于《国际交互设计与制造期刊》，2017年。[23] M. Poret、S. Ibarboure、M. Robine、E. Duc、N. Couture、M. Desainte-Catherine 和 C. Semal，《用于3D打印过程监控的声化》，发表于*2021*年国际计算机音乐会议论文集（*ICMC 2021*），智利圣地亚哥，2021年。[24] C. Xia、Z. Pan、J. Polden、H. Li、Y. Xu、S. Chen 和 Y. Zhang，《电弧丝材增材制造综述：监控、控制与自动化系统框架》，《制造系统学报》，第57卷，第31–45页，2020年。 [在线]。获取地址：https://www.sciencedirect.com/science/article/ pii/S0278612520301412[25] W. W. Gaver，《我们听见了什么？听觉事件感知的生态学方法》，生态心理学，第5卷第1期，第1–29页，1993年。[26] M. M. Blattner、D. A. Sumikawa 和 R. M. Greenberg，《耳标与图标：其结构与通用设计原则》，人机交互，第4卷第1期，第11–44页，1989年。[27] W. W. Gaver，《SonicFinder：使用听觉图标的界面》，人机交互，第4卷第1期，第67–94页，1989年。[28] N. A. Macmillan 和 C. D. Creelman，《检测理论：用户指南》。心理学出版社，2004年。