

日本交流项目汇报

2018. 9. 18

胡昕悦 201728060049
王鹤燠 201728060150
王爽 201728060156
务孟瑶 201728060168
杨璐溪 201728060185

有幸到日本高知工科大学任向实老师的人机交互实验室交流学习 4 个月，我们 5 个人收获良多。在实验室中我们每个人都有确定的负责人带领，每周会有固定时间跟负责人开会，也会定期跟任老师汇报我们的项目进展。由于大家初期对于科研项目并不很了解，需要不断地跟项目负责人探讨问题，交流彼此的想法。做科研需要持之以恒，不畏困难的态度，更要对所做的工作怀有热情。下面就是我们在交流期间各自的项目简介。

手机点击行为下的手眼协调观测

胡昕悦 201728060049

选题背景

手机目前已经成为现代生活不可或缺的一部分，我们几乎每天都在使用智能手机。显然，对于目前的智能手机来说，点击操作是所有手机交互操作中最常用的交互方式。在人际交互领域，对于点击操作的研究已经有很多，但是到目前为止，还没有对手机点击行为下的手眼协调的研究。手眼协调是指人在视觉配合下收的精细动作的协调性，也可以定义为是我们的眼睛（视觉）和手（运动）的能力。手眼协调是一个重要的，复杂的认知能力：

- 我们用眼睛来引导注意刺激，帮助大脑了解身体在空间中的位置（自我感知）。
- 我们用双手同时进行我们的眼睛接收的视觉信息的任务。

对于手机点击操作这种精细的动作来说，手眼协调扮演着重要的角色。研究清楚点击操作下的手眼协调模式之后，可以作为以后界面交互的设计指导。因此，我觉得选取这一问题进行研究。

文献调研过程

在确定了研究主题之后，我开始着重研读手眼协调与点击行为领域的相关文献。在手眼协调领域，我先从宽广的概念出发去了解相关研究。

手眼协调在我们的日常生活中占有重要位置。例如，当我们需要拿起茶杯喝水，或者想要抓住飞过来的球的时候，都需要手眼协调工作。已经有许多研究者在研究人们操作物体时候的手眼协调行为。Roland S. Johansson 与其同事做了一个关于人们在操作物体时候，凝

视行为与手指之间运动协调。他们发现，当人们想要抓住一个物体并将其移动到目标位置时候，眼睛的凝视会支持手的运动，并且会标注一些关键性的位置，以便于之后手的移动。换一种方式来说，即是眼睛的凝视会稍微领先于手的运动，并且会为手的运动做一些标记。

【Johansson R S, 2001】Emma Gowen 分析了人们的在 Drawing 和 Tracing 简单图形时候，两种不同操作下的手眼协调行为差异。他们的研究结果呈现：手眼行为的耦合度在 Tracing 的时候比在 Drawing 的时候更加耦合，这可能是因为在进行 Drawing 和 Tracing 的时候，会使用不同的策略：Tracing 需要精细的对比，需要更多的视觉反馈【Gowen E, Miall R C, 2006】。

关于点击任务下的手眼协调行为，也已经有一些研究成果。例如：Barton A. Smith 和他的同事观测了在电脑上的点击任务下的手眼行为表现。他们发现了在电脑屏幕上操作点击任务时的三种手眼协调策略，分别是：（1）眼睛行为领先于光标点击行为；（2）眼睛跟随光标运动；（3）眼睛在光标和目标之间切换【Barton A. Smith, 2000】。

在进行了大量相关文献的研读之后，我确定了自己可以切入的角度。目前关于点击行为手眼协调的研究都集中于手和光标的协调以及手和电脑屏幕上光标的协调，还未有关于手和手指直接点击位置协调的研究，并且是在手机这样的相对较小的屏幕上的研究。因此，这更加坚定了我的研究主题的意义。

实验设计

本研究属于比较基础的行为观测的研究。因此，主要采用基础实验设计的心理学研究方法。实验需要观测的是被试手和眼睛在操作点击时候的协调行为，因此，需要分别测量手的实验数据以及眼睛的实验数据。

本实验使用的实验程序，是基于安卓系统编写的一套基本的点击操作实验。使用的实验手机型号是三星 Galaxy S6。

本实验中，目前手的实验数据提取有两种方案，一种是使用 VICON（动作捕捉系统）设备来记录；另一种是使用摄像机记录手的动作，之后根据一套算法来定位手指位置；而眼睛的数据则是需要使用眼动追踪仪器来记录。

实验搭建过程中的难点就是将两套设备连接互通同时开始记录手眼数据。

本实验属于基础的行为观测实验，本次试验共需要 12 名被试。被试在执行实验过程中，需要头戴 Dikablis 眼动仪，右手食指粘贴 marker，之后在手机上执行三个不同的点击实验。实验的顺序是 1D 任务，2D 任务和搜索点击任务。

实验一：基本的 1D 点击任务

这是一个经典的 Fitts Law 任务。打开手机应用，填写实验编号之后，选择实验一。每次试验开始之前，手机屏幕中央会出现一个 96px 宽的灰色小条，作为试验开始的位置（被称作 starting bar）；每次点击灰色的 starting bar 之后，出现红色的 target bar，要求实验参与者以既快又准的方式点击红色的 target bar。Starting bar 的宽度和其中心点到 target bar 的中心点的距离在每次试验中按照预设的条件组合随机显示给实验参与者。要求被试使用食指完成点击任务，每次点击之后，食指回到灰色 starting bar 的位置。每个实验控制变量组合各重复 5 次。如图 1 所示：

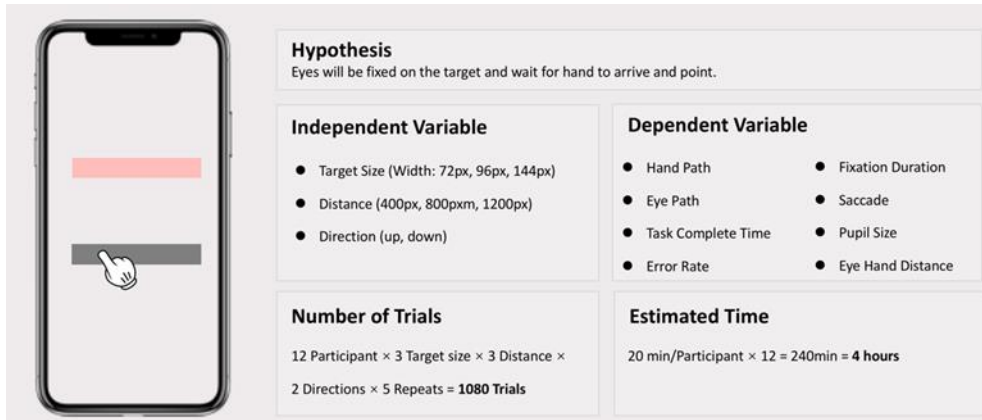


图 1 - 实验 1

此实验包含三个控制变量D,W 和 Direction。其中D代表 starting bar 的中心到 target bar 中心的距离；W 代表 target bar 在手指运动方向上的宽度（本实验为 target bar 在上下方向上的宽度）；Direction 代表 target bar 相对于 starting bar 的方向。

D 的取值包含 400px, 800px 和 1200px；W 的取值包含 72px, 96px 和 144px；D 的取值包含南和北。以上参数参考了三星 S6 手机的界面元素尺寸。

本次实验共包含 12 被试 * 3D * 3W * 2 Direction * 5 重复 = 1080 次试验。

实验二：基本的 2D 点击任务

点开手机应用，填写实验编号之后，选择实验二。每次试验开始之前，手机屏幕正中央会出现一个直径为 96px 的灰色圆圈，作为试验开始的位置(被称作 starting point)；每次点击 starting point 之后，系统给出当前试验的红色圆形 target，要求实验参与者以既快又准的方式点击红色的 target bar。Target 的直径和其圆心到 starting point 的圆心的距离在每次试验中按照预设的条件组合随机显示给实验参与者。被试被要求使用食指完成点击任务，每次点击之后，食指需回到灰色 starting point 的位置。每个实验控制变量组合各重复 5 次。如图 2 所示：

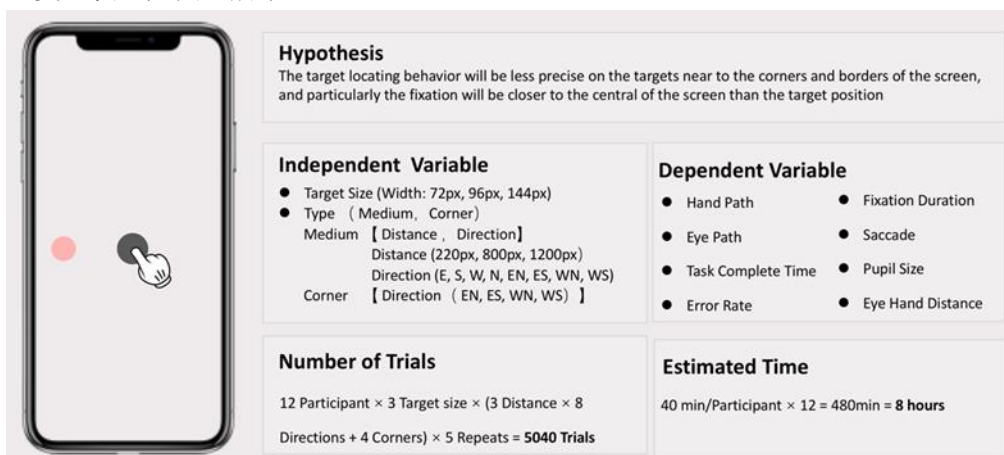


图 2 - 实验 2

此实验同样包含两个控制变量W 和 Type。W 代表圆形 target 的直径；W 的取值包含 72px, 96px 和 144px；D 的取值包含是东，南，西，北，东北，南北，西南，东西。Type 代表目标出现的区域。目标区域的 Type 的取值有两个，分别是：Center 和 Corner。在 Center 的情况之下，控制目标出现的变量有两个分别是：D(Distance)和 direction。其中D代表 starting

point 的圆心到 target 的圆心的距离, D 的取值包含 220Px, 440px, 660px; direction 代表 target 相对于 starting point 的方向; 在 Corner 的情境之下, 目标出现的位置是手机屏幕上的四个角落: 左上角, 右上角, 左下角和右下角。角落位置至 start point 圆心的距离为 1400px。以上参数参考了三星 S6 手机的界面元素尺寸。

本次实验共包含 12 被试 * 3W * [(3D * 8 directions) + 4 个角落位置] * 5 重复 = 5040 次试验。

实验三: 搜索点击任务

为了更好地研究人们使用手机进行点击任务时的手眼协调行为, 我们在 2D 点击的基础上增加了视觉搜索任务。

点开手机应用, 填写实验编号之后, 选择实验三。每次试验开始之前, 手机屏幕正中央会出现一个直径为 72px 的灰色圆圈, 作为试验开始的位置(被称作 starting point)并在圆圈上方显示 target 的数字提示; 要求被试确定看清楚并记住该数字之后, 点击圆形 starting point 开始试验。试验开始时, 屏幕上显示 (3D * 8 directions) + 4 个角落位置 = 28 个内含数字(两位数)的圆圈。要求被试使用食指点击包含与刚才记忆中数字相同的目标。每次试验中整体的布局不变, 但圆圈内数字随机改变。每个屏幕位置上的点击重复 5 次。如下图 3 所示:

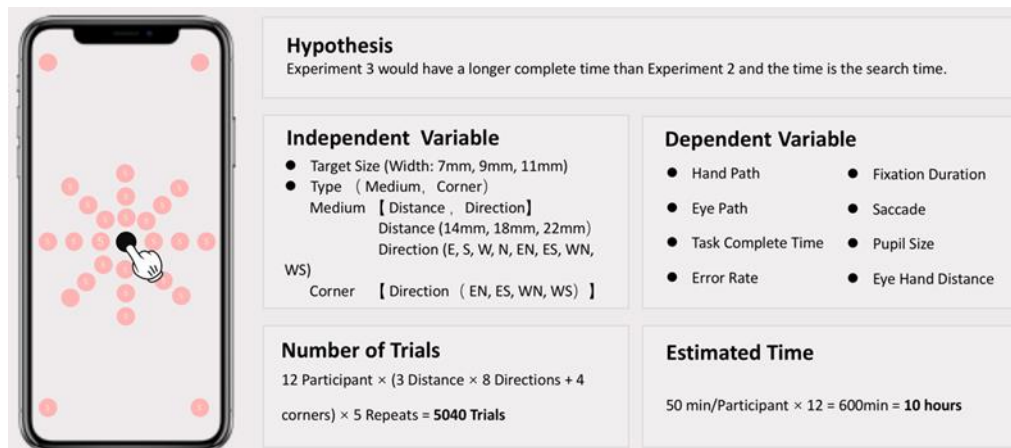


图 3- 实验 3

此实验的控制变量为 D 和 Direction。其中 D 代表 starting point 的圆心到 target 的圆心的距离; Direction 代表 target 相对于 starting point 的方向。

D 的取值包含 220Px, 440px, 660px; Direction 的取值包含东, 南, 西, 北, 东北, 南北, 西南, 东西, 另外 target 的位置包含手机屏幕的左上角, 右上角, 左下角和右下角, 角落位置至 start point 圆心的距离为 1400px。

本次实验共包含 12 被试 * [(3D * 8 directions) + 4 个角落位置] * 5 重复 = 5040 次试验。

实验设计经过多次的小组讨论, 已经趋于完善。

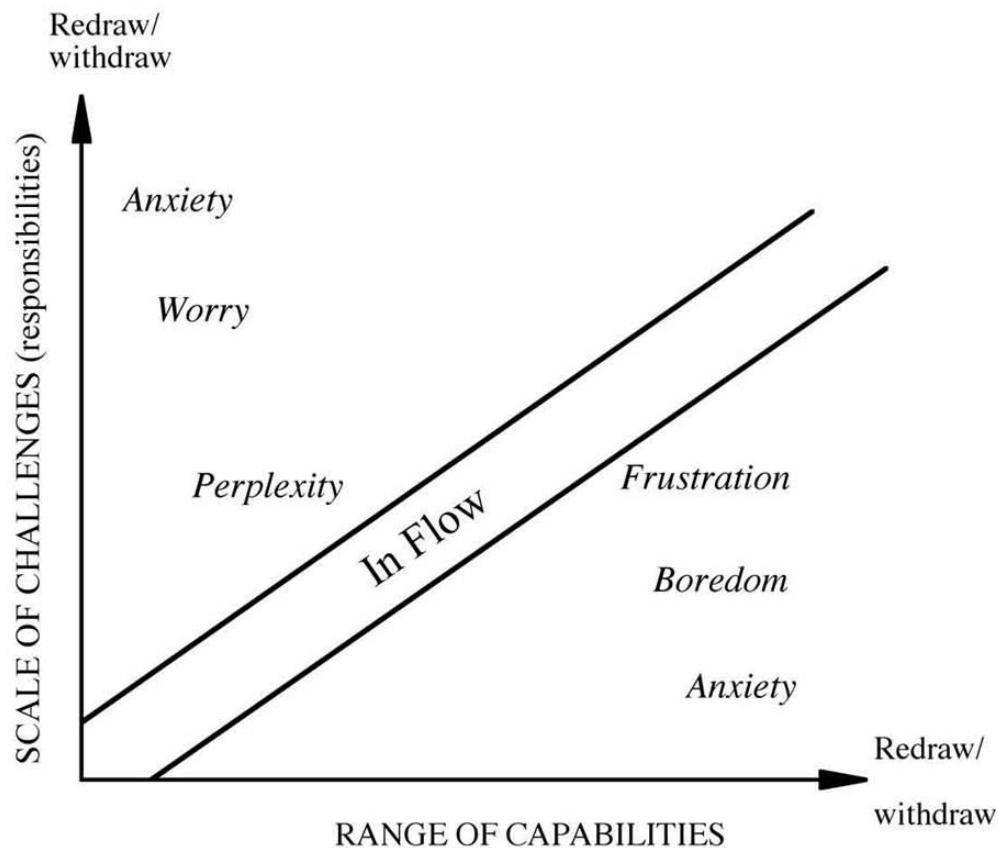
下一步便是进行被试的招募与实验的实施。

视频游戏中实验诱导的心流体验与眼动追踪的相关性研究

王鹤燠 201728060150

项目背景

心流（Flow）体验的概念于 1975 年由心理学家米哈利·希斯赞特米哈伊提出，指一种个体在专注进行某行为时所表现的心理状态（见图一）。心流体验发生在日常生活中的很多情境中，有研究者便指出，米哈利的心流概念正是描述主观游戏体验的最突出的理论之一（Klasen, 2012）。目前测量心流体验最常用的方式为问卷与访谈，但是心流体验出现在活动过程中，而在填写问卷和接受访谈时，被试却已脱离心流状态，因此自我报告的测量方法具有极大的主观性，且得到的结果并非实时的反馈。



Adapted from Csikszentmihalyi, M. *Optimal Experience*
Cambridge University Press 1988

图一 心流（Flow）图示

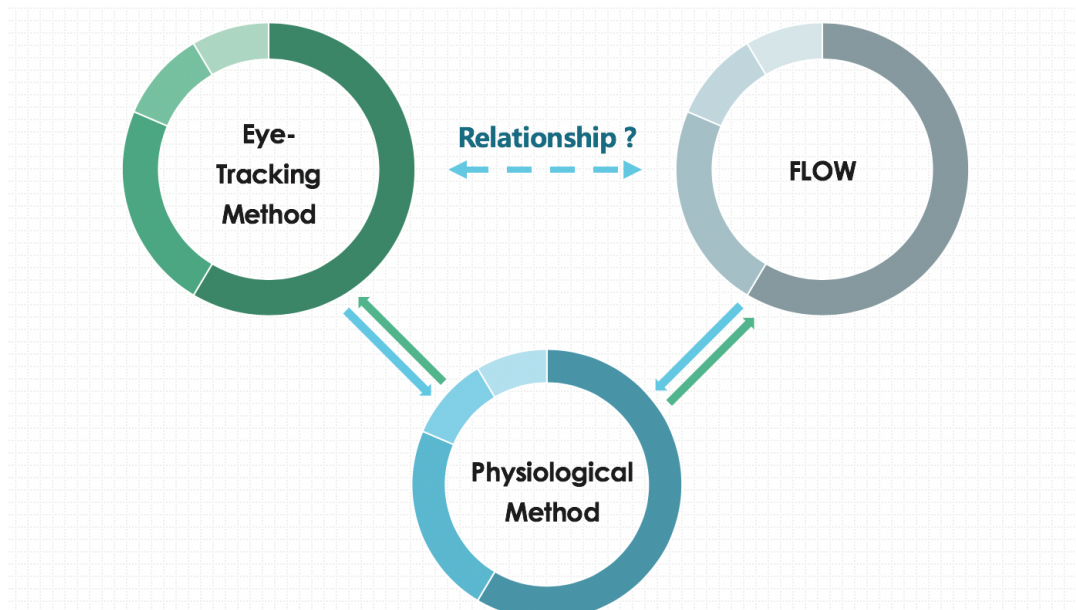
目前已经有一些研究人员进行了生理指标与心流状态的相关性研究，研究结果涉及心流状态下脑电、皮肤电、呼吸、唾液、心血管等生理指标的特征。但是将生理指标作为心流体

验的测量手段有一大弊端在于生理传感器穿戴复杂，被测量者无法进行大幅动作的活动。同时，作为突出描述主观游戏体验的心流体验，在一定程度上可被视为衡量游戏体验好坏的标准，甚至可以通过探知玩家的情绪状态，自适应调节游戏难度，使玩家在心流状态下获得更完美的游戏体验。而复杂的生理传感器设备无疑将对游戏体验带来负面影响，在 AR、VR 及体感游戏日渐盛行的今天，生理指标测量心流状态的方法已不可取。

研究问题

视频游戏在娱乐，学习和健康方面发挥着重要作用。从游戏行业诞生开始，设计师就一直试图让游戏玩法更具吸引力并且在更长的时间内具有挑战性，以改善玩家体验并避免挫败感。最常见的方法是技能平衡，它根据玩家的技能自适应地改变游戏难度。然而，很少有人研究过情感适应（即，根据玩家心理状态改变游戏机制），这可能是改善玩家体验的潜在且更有效的方法。这项工作旨在系统地研究基于眼睛的情感适应对玩家体验的效果。

然而，纵使有大量的相关研究，关于眼动指标和玩家体验之间的关系依然知之甚少。在这个项目中，我们试图回答以下问题：在游戏过程中哪些眼动指标最能代表情绪状态？如何通过眼动指标来区分基本的玩家体验，如冷漠，无聊，心流和焦虑？是否可以通过眼动指标识别不同的挫折感，例如，任务与技能不平衡、模棱两可和缺失反馈？



图二 实验目的

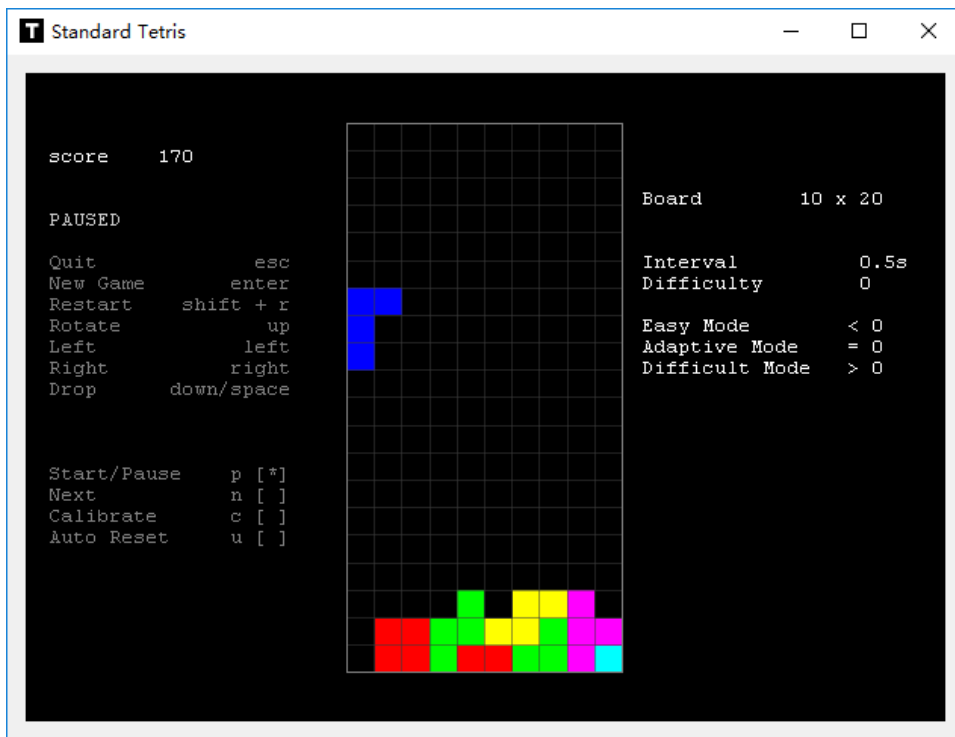
实验设计

在确定了具体的研究方向与题目后，项目便进入了具体设计实验细节的过程。首先需要确定的是实验中需要测量的眼动与生理数据，即因变量。在阅读总结了大量相关文献后，我同 Mentor 讨论选择了皮肤电指标、呼吸指标和心率变异性指标作为衡量心流状态的客观指标以及帮助我们关联眼动与心流的中介指标。

确定了实验因变量后，实验的自变量就成为了关键。由题所知，该实验的实施环境为视频游戏，因此选择什么游戏用于实验也至关重要。于是我们寻找了涉及到心流体验与游戏的相关文献数十篇，在仔细研读与总结后，选择了为多数研究所使用的俄罗斯方块，同时，也

从相关文献中获得了游戏难度的具体控制手段，即俄罗斯方块游戏中方块的下落速度。

确定了用于实验的游戏后，接下来便是编程实现游戏使其能够应对符合实验要求，即实现游戏的自适应难度调节，以使得被试在自适应模式的游戏过程中迅速进入心流状态。别看这实验要求如今写来只是短短的一句话，可是这个问题的解决却足足耗费了我一个半月的时间之久。尽管在之前我听从 Mentor 的指示在国内自学了 C# 编程语言，可是短短几个月见缝插针的学习，在面临真正待解决的问题时，仿若杯水车薪。好在俄罗斯方块在网络上有着大量的开源代码，我选择下载了数个不同版本，在反复地构建、调试与修改后，终于基于其中一种较为完善的源代码搭建完成了最终实验用的带自适应难度的俄罗斯方块游戏。尽管游戏中仍有许多不足之处亟待改善，但是对于作为编程门外汉的我而言，已经是一个令人欢呼雀跃的成果。



图三 俄罗斯方块游戏界面

实验的硬件已准备好，在正式进行实验之前，实验的每一个细节都需要仔细考量，于是 Mentor 要求我参照往年实验室发表在 CHI(人机交互顶会)上的文章，进行实验设计的拟定，并在组会上向研究所所有成员发表，以基于其他成员的各种反馈，进行进一步的修改完善。最终我们的实验设计如下：

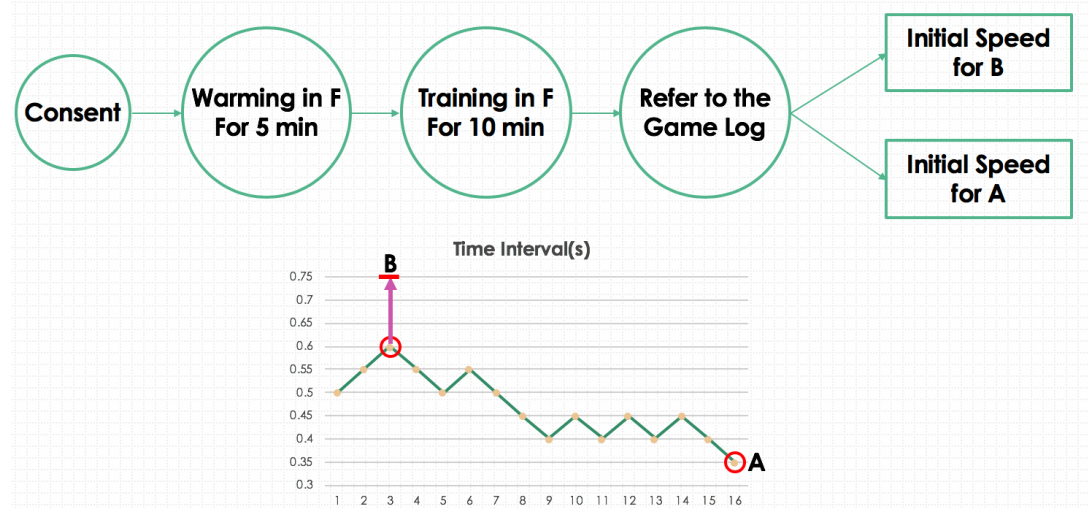
由于随着方块下落速度的增加，游戏难度也随之增加，我们设计了三种游戏模式，通过控制速度来调整难度，即简单模式，自适应模式和困难模式。

在简单模式下，无论玩家清除了多少行，方块都会以非常慢的速度下落，此模式对应无聊情境（此后记为 B）。在自适应模式中，方块的下落速度适应玩家的表现。根据测试实验，我们在这种模式下决定了速度改变规则：如果玩家使用不超过 15 个方块成功消除了五行及以上，则难度会自动增加一级。如果玩家已经使用了 15 个方块但只消除了三行或更少，则难度降低一级。自适应模式对应心流情境（此后记为 F）。在困难模式中，方块最初便以一个非常快的速度下降，且每消除十行，难度增加一级，这个模式对应焦虑情境（此后记为 A）。

实验开始前，被试在签署同意书并了解实验目标后，首先将在自适应模式下玩 10 分钟游戏作为训练。根据 10 分钟游戏中被试的表现，我们确定了简易模式和困难模式的初始速

度：由于方块的下落速度会随着玩家的表现而变化，在 10 分钟的游戏过程中，必然会有一个最快速度和一个最慢速度。我们通过游戏数据日志找出这两个速度，然后将简单模式的初始速度设置为比最慢速度低 3 个等级；而困难模式的初始速度设置为最快速度。

Task & Procedure



图四 实验前热身与训练流程

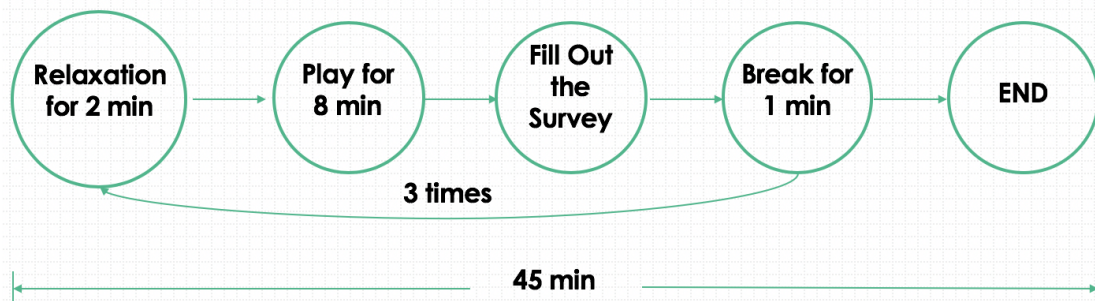
正式实验部分，被试将被分成 3 组，每组 5 名被试。每组被试将以不同的难度模式顺序玩 6 次俄罗斯方块，顺序如下：

- 1、B-F-A-B-A-F
- 2、A-B-F-A-F-B
- 3、F-A-B-F-B-A

被试每次玩 5 分钟，中间休息 2 分钟，同时填写心流的测量量表。整个实验过程将持续大约 90 分钟。

在正式实验前的预实验中，Mentor 作为我的被试完成了以上全部实验过程，但是我们发现了许多之前未能考量到的纰漏。于是我们又对实验步骤进行了又一次的修改与完善：要求相同性别的被试参与实验；十分钟的训练之前，再增加五分钟的热身，以帮助被试熟悉游戏操作；困难模式的初始速度为十分钟训练中的最快速度，而不再提升一个等级；每名被试将玩 3 次俄罗斯方块，每次 8 分钟，在每次玩游戏之前，还将观看 2 分钟的放松视频，以获取被试生理、眼动指标的基准线，有助于实验后的数据分析。同时还设置了被试背景调查问卷，和心流倾向调查问卷，要求被试在参与实验之前进行填写，并根据其填写内容，为其分组，以确保每组被试间游戏水平相当。

Task & Procedure



图五 更改后的正式实验流程

将游戏次数由六次改为三次，考虑到实验的平衡性，我们应进行 B-F-A、B-A-F、A-B-F、A-F-B、F-A-B 及 F-B-A 六种顺序的实验，而每组实验应收集至少两名被试的数据，即共需做至少 12 次实验，然而，由于我在日本的时间有限，余下的时间不足以让我做完所有实验，因此，我们选择了 B-F-A、A-B-F 和 F-A-B 三种顺序，每组实验收集三名被试的数据，共计 9 次实验。收集到的数据由我回国后进行分析确认实验假设是否可行。并根据数据分析结果来决定余下三种顺序的实验是否需要 Mentor 再招募九名被试实施并完成。



图六 实验实景

在对 9 名被试实施实验的过程中，出现过硬件上的失误，也存在过与被试直接语言沟通的障碍，但是总体而言实施效果较为理想，成功收集到了 9 名被试一次训练、三次放松、三次游戏共计七种情境下的生理与眼动数据。而实验实施的最终效果，则需看最终数据分析的结果。

目前该项目实验的数据分析正在顺利进行中，希望最终能得到合理有效的结果。

触屏设备上的文本修改技术研究

王爽 201728060156

背景

文字是用户与设备交互，实现多用户之间沟通的一种重要形式，文字的输入过程更是人机交互过程中非常重要的一环。在普遍情况下，用户需要保证其输入或处理的文字内容没有错误（无论是从字符层面，还是语意层面）。在文字输入的过程中，往往有多种因素会导致输入的文本出现错误。以智能手机为例，用户用食指和（或）拇指在屏幕中显示的虚拟按钮上点击，以完成字符的选择和输入。与传统的物理键盘不同，触摸屏无法提供足够的触觉反馈，与此同时，手指在点击虚拟按钮时对按钮产生遮挡，导致无法准确评估点击位置，产生“胖手指”问题；除此以外，由于屏幕尺寸的限制，每一个虚拟键盘按钮的尺寸都比较小，不仅加重的“胖手指”问题，而且增大了点击的难度。从识别的角度讲，由于技术发展的限制，在用户不是非常精准的点击之后，机器的识别往往也不能完全精确。还有用户本身对单词、汉字的拼写也会导致出现各种错误。尽管有一些辅助输入的方法，如预测列表等，但只能在一定程度上减少错误。另外，在使用一些技术完成大量文字输入过程中，如 OCR、语音识别等，由于算法和识别精度的限制，往往会有大量错误出现。如此种种都可以看出在文本输入的过程中，文本错误仍然无法完全避免。那么当错误出现时，用户则需要对此做出响应，即展开错误修正工作。

研究问题

在文本出现的错误中，有很多能够在输入过程中及时发现错误并通过退格键或者辅助的预测列表来修正。当然也有被忽略后来才发现的，比如有的用户会习惯打完一句话后读一遍再发送消息，发现错误再修改，这种错误有的发生在句尾，容易通过退格键完成删除修改。但如果是发生在句子中间或者靠前部位，手指移动光标就有一些问题，整体修改下来会很耗时，给用户及其不好的体验。

因此本课题着眼于触屏设备上的文本错误修改问题，尤其是出现在文本中间或者靠前部位被忽略的错误（插入、删除、替换）的修改策略以及实现。通过分析现有的研究以及产品中的解决方案找到可以改进的地方进行深入讨论，并且尝试关注文本错误修改的整个流程而非具体的某一方面，在流程上进行优化调整，并且针对调整后的流程给出实现方案，然后进

行与传统流程指导下的传统修改方法进行实验比较，验证新流程以及方案的有效性。为智能设备上文本错误的修改提供新的角度和解决方案。

研究过程

在确定了选题之后，开始了文献的查找、阅读和整理的工作，查找到的相关文献主要可以分为三大类：文字输入技术；文本修改；交互行为研究。通过阅读文献，了解到文本错误产生的原因，主要就是上面提到过的用户本身的拼写问题，“胖手指”问题，以及技术发展的限制。同时，也对当前市面上的产品在面临这个问题是如何解决的做了调研，收集到包括百度、苹果、三星、谷歌等虚拟键盘上的解决方案，以及一些专门做键盘的第三方给出的方案（如 flekxy），前人的研究成果以及以后的解决方案是我们进行进一步研究的基石，经过整理分析得出一个现有的改错流程图（如图所示）。

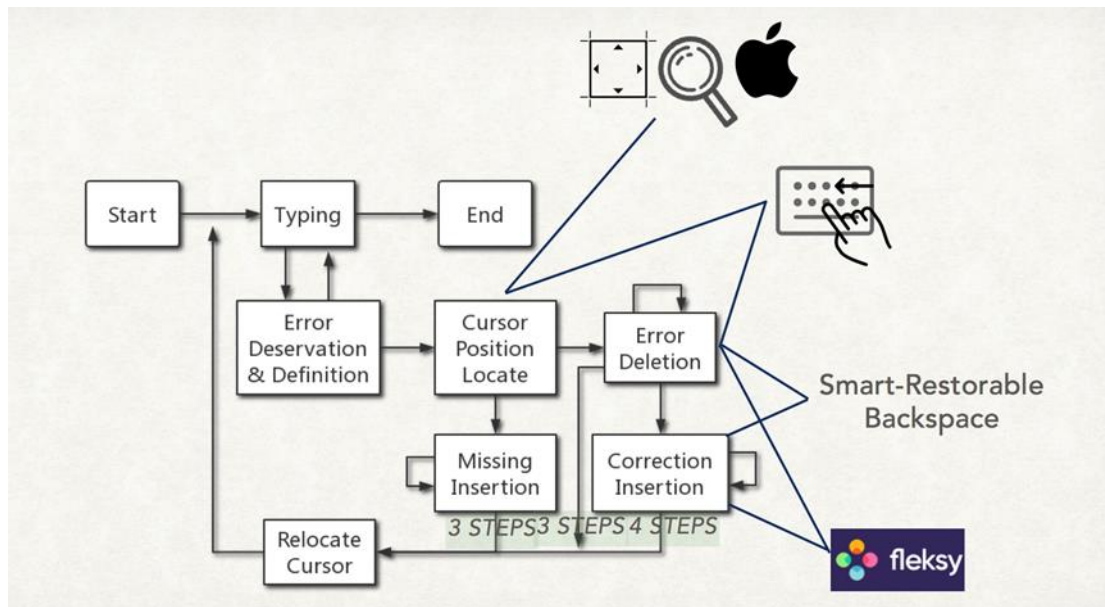


图 1 文本错误修改流程及不同解决方案的侧重点

现有解决方案虽然侧重点各有不同，但从整体来看都遵循同样的流程。即：用户在输入过程中发现靠前部位的错误同时已经判断出错误的类型，然后将光标移动到错误发生的位置，接下来根据错误类型决定是增加还是删除或者删除后增加的动作，完成修改后在将光标移动到句尾继续输入。一般来说，产品中的解决方案更加关注如何让定位更加容易，毕竟这也是整个体验中最糟糕的部分，当然也有将其中的一两个步骤结合起来解决问题的。事实上，我个人认为苹果的 3Dtouch 解决方案在定位上已经将体验改进的不错了，但其也是在这个相同的流程中在“定位”这一个步骤发力。经过讨论研究，我们认为此流程有可改进的可能，因而我们打算在整体流程上对改错过程加以改进，并在此新的范式的基础上提出新的设计方案。

在设计新的改错方法的过程中，我和研究室成员从日常生活以及类似的场景出发，试图寻找最自然的方式，提出了几个设计设想（图 2，图 3）。把这些设计构想都整理起来，与研究室成员进行讨论评估，最终确定选择的设想是从射击游戏以及泡泡龙游戏中获得启发提出的最初的设计雏形，其初衷是将游戏化的概念引入文本错误修改的过程中，消灭改错过程中的消极情绪，也就是将正确的字符当作炮弹由用户操控发射到正确的位置，就算有定位不准

的问题，但在这个过程中充满趣味，可以在一定程度上改善体验。

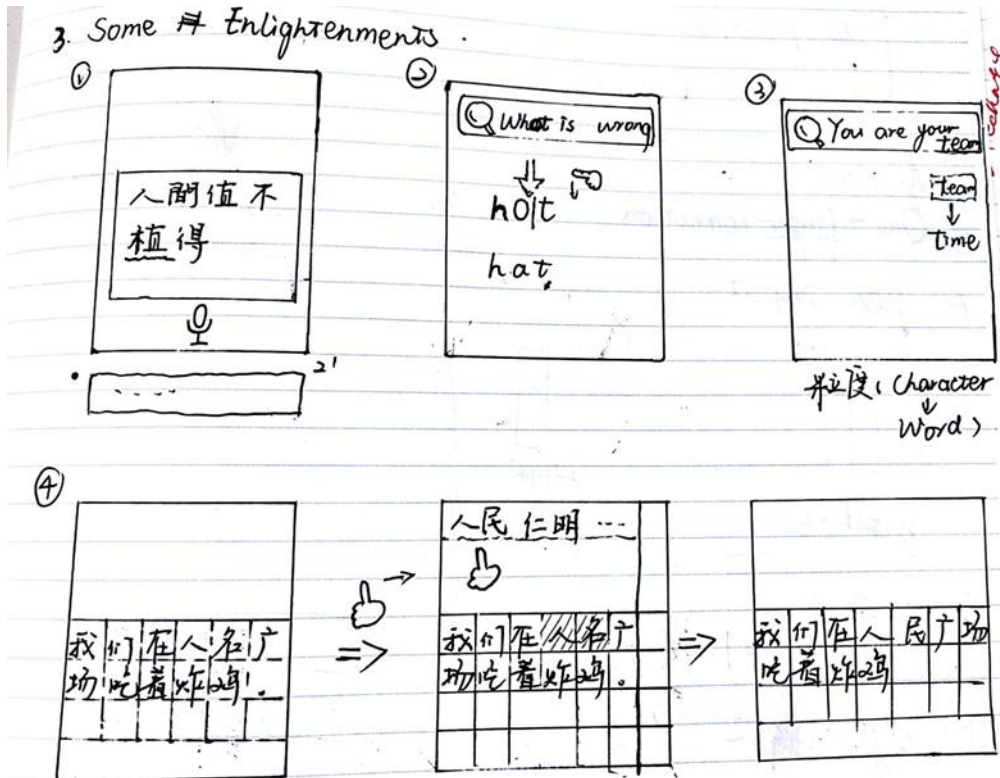


图2 触屏设备上的文本错误修改设计方案 1-4 草图

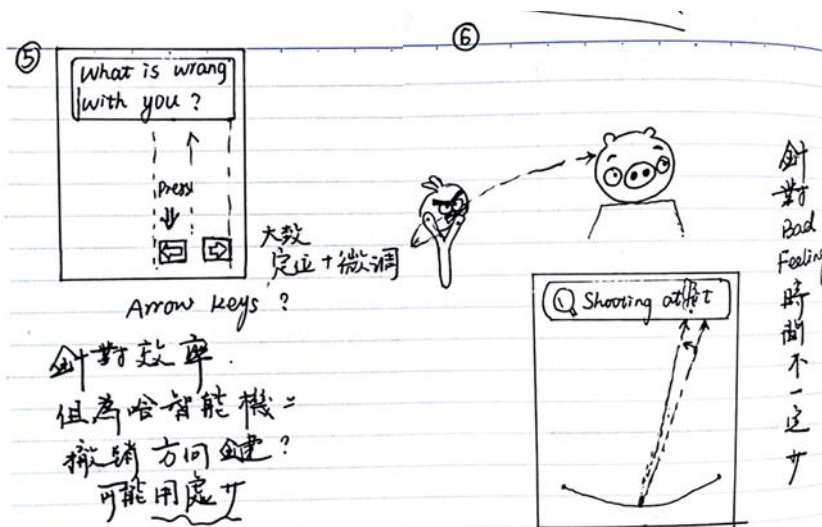


图3 触屏设备上的文本错误修改设计方案 5-6 草图

经过多轮讨论，互相激发产生新的想法，对方案6进行改进提升并然后讨论可行性，进行内部评估等，最终确定了设计方案。这个方案从游戏泡泡龙中得到启发，描述起来就是在用户发现错误并判断了错误类型之后，第一步不是去移动光标，而是先呈现出解决方案（正确的字符，触发改错模式后，由用户自行输入），然后就是“发射”的过程，将正确的字符送到正确的位置（后台进行字符串匹配确定大概位置），实现修改（用户在需要修改的大概位置触摸即可）。目前的技术从策略层面讲删掉了定位光标 and 回移光标的步骤，如果是针对替换类型的错误，将先删除再增加的过程结合到一起，总体上来讲，更加简洁直接。

接下来要验证新设计的表现，需要进行对比实验，目前仅进行到实验设计的阶段。因为之前没有过学术实验的经历，需要学习很多实验设计相关的知识，从文献中学习类似实验的考虑因素、设计过程和实施过程，在博士师兄的指导下尝试设计实验，讨论修改，可能在实验环境搭建结束甚至进行实验前仍然会对实验进行修改完善。由于时间的限制和自身编程知识储备不足，实验平台的搭建工作由博士师兄进行，实验环境尚未搭建完成，因此在此期间未能完成实验，后续会继续跟进。

应用文本输入评估方法的 2D 点击技术的评估研究

务孟瑶 201728060168

背景

这些年来点击选择技术已被广泛研究。为了比较它们之间的优劣，评估方法通常考虑它们的时间表现，错误率和基于 Fitts 定律的数学模型。众所周知，对于文本输入方面的指标也有很多评估，如 WPM, KSPC, ER 等。因此，文本输入的相关评价也是一个备受推崇的话题。通过研究对比发现，其实在文本输入方法的评估指标中，有相当一部分可以帮助评估点击技术。所以我们的项目目标就是将评估文本输入的指标转化为评估点击技术的指标。

在过去几十年中很多学者已经广泛研究了指向-选择-点击技术。实际上，选择目标的对象是启动进一步交互的最基本任务之一。此外，点击选择技术在不同的情形下差别很大，例如桌面环境，远程屏幕，移动环境，增强现实环境，虚拟现实环境等。除此之外，点击选择研究涵盖了广泛的输入特征，像直接或间接输入模态，例如手指触摸，手指指向，头的旋转，用眼睛凝视等等。由此可见指向-选择-点击技术的复杂性。

新的选择点击技术或设备的评估通常通过其关于选择时间（选择目标的平均时间）和错误率（错误选择的平均数量）的性能来完成。除了这两个定量测量，HCI 研究人员还经常提供遵循或扩展 Fitts 定律的研究结果协议。事实上，Fitts 定律使 HCI 研究人员能够有章可循地去做研究，而且近些年已有相关学者对于 Fitts 定律在 HCI 领域地应用呼吁使用相同地标准，以便于各个研究项目间的比较。

文本输入是人机交互领域中另一个历史悠久且活跃的话题。文本输入技术使用户能够通过选择字符序列来输入文本。相对于指点选择技术，文本输入技术涵盖也同样涵盖了广泛的使用情境和输入特征，其复杂性不亚于点击选择技术。

研究问题

然而，与点击选择技术相比，文本输入技术受益于更多的定量性能指标，其评估项目和指标相对更为完善。这使研究人员和设计人员能够更好地了解特定技术的潜在影响。例如，总误差率表示文本输入期间和之后的准确度，允许考虑纠错成本，这个条件的描述就很是详细了。

所以,我们开始这项工作的动机就是点击选择技术和文本输入研究领域之间的密切关系。首先,点击选择技术可以被视为文本输入技术的核心组成部分。虽然前者通常用于选择目标,但后者用于选择目标字符序列。其次,我们可以在它们的不同使用环境(例如,桌面或虚拟环境)和它们对输入特性的覆盖(例如,直接,间接,手指触摸或凝视)之间定义明确的并行。鉴于这种接近程度,我们的目标是为量化评估指标提供另一个平行线。我们希望通过相对已经比较完善的评估文本输入方法的指标进行分析,将其转化为对于点击选择技术的评估上面,并且为之后的点击选择技术的评估方法的研究提供一定的佐证。

研究过程

通过大量的文献回顾,我们了解到目前为止,我们知道很多方法可以进行“文本输入”。无论内容是由 QWERTY 键盘还是 12 键键盘输入,还是微型设备上的各种新键盘或方法,我们都需要一些标准指标来评估它们。许多研究人员研究文本输入的优化或创新之所以类似,是为了提高速度,降低错误率和提高输入的便利性。





Technique	Approach	Execution	Size(mm)	QWERTY	Publisher &Year	Phrase set	Performance	Image
ZoomBoard	Zoom	Two step tap		Y	CHI 2013			
SwipeBoard	Gesture	Two step swipe		Y	UIST 2014			
ZShift	Pointer shift	Finger drag and release		Y	CHI 2015			
Callout	Pointer shift	Finger drag and release		Y	CHI 2015			

图 1 部分文本输入方法研究总结

点击技术的评估与文本输入更相似。事实上,点击技术可以被视为一个特殊的“文本输入”,只需一次点击即可。当然,Fitts 定律被认为是点击技术评估的基础。

在回顾了一系列的文本输入和点击选择技术的评估方法之后,我们决定首先详细描述每个文本输入度量,并在应用于点击选择技术时提供理论解释。然后,我们再设计实验评估熟知和已建立的点击技术。

图 2

这几个月以来,我们项目组研究了有关现有文本输入和点击技术的评估指标。我们想通过比较它们之间的相同和差异,进而探讨之间相关联的应用方式。

对于文本输入部分,有一些常见的指标,如文本输入速度(WPM),错误率(ER),效率(KSPC)。除此之外,许多研究人员根据一些基本指标提出了自己的评估指标,所以有非常多的测量指标的公式及其变形。因此,很难比较研究或从这一工作中提取有意义的平均文本输入速度和错误率作为对比。考虑到现行实验采用指标的计算方法的不一致性,我们决定使用 Soukoreff 和 MacKenzie 引入的指标。研究项目中对指标变换的解释,可以视为这些指标和公式的变形。

对于点击选择技术部分,有一些类似的指标,包括文本输入的度量,如错误率。当我们

提到点击技术时，人们会想到 Fitts 定律，有关学者也已经在 2004 年提出在人机交互领域供参考的 Fitts 定律的使用标准。由于 2D 和 3D 点击选择任务之间存在很多差异，因此 Fitts 定律的应用也不同。我们发现了一些关于不同尺寸 Fitts 定律（1D，2D 和 3D）的计算研究。

正如我们之前所说，我们认为文本输入和点击技术之间存在一些共同点，因此我们尝试将一些文本输入度量应用于点击技术。当然，由于测量指标所要求的详细说明与解释，以下列举一些转型的结果：

		Text-entry metrics		Pointing techniques evaluation	
		formula	interpret	conclusion	formula
P	Presented Text		what participants had to enter, and P is the length of P.		the number of targets that participants has to select
T	Transcribed Text		the final text entered by the participant, and T is the length of T		the final number of selected targets by the participants
IS	Input Stream		the text that contains all keystrokes performed while entering the presented text, and IS is the length of IS.		the series of operations that contains the correct operations performed while selecting the showed targets
C	Correct	$C = \max(P , T) - MSD(P, T)$	the number of correct characters in the transcribed text.		the number of correct clicks in the final chosen targets (selected target)
INF	Incorrect Not Fixed	$INF = MSD(P, T)$	the number of unnoticed errors (incorrect characters) in the transcribed text.		the number of missing targets in the final chosen targets (selected target)
F	Fixes		keystrokes in the input stream, which are edit functions (backspace, delete, cursor movement, etc), modifier keys (shift, alt, control, etc.), or navigation keys (left, right, mouse click, etc).		total operations (move-click-move-click)
IF	Incorrect Fixed		keystrokes are those in the input stream that are not editing keys (F).		the number of first click which is incorrect (unselected target)

图 2 部分指标转换公式及解释

- 目标：目标出现在屏幕上，参与者需要通过鼠标点击目标。
- 选定目标：当目标出现时，参与者使用鼠标选择它们然后单击。但也许并非每次点击都可以获得目标，有时参与者需要更多点击才能获得目标。因此，选定目标意味着所选目标的最终数量。
- 操作流：与文本输入指标一样，输入流表示在输入时执行的所有击键文本，包括“退格，左，右……”当参与者使用鼠标完成指点选择任务时，他们将进行一系列操作，如“移动 - 点击 - 移动 - 点击”。因此，操作流意味着当参与者选择显示的目标时执行的正确操作。
- 正确
 1. 第一次正确点击：每个目标的第一次点击是正确的。我们的想法是获得第一次点击的成功率，因为在实际情况下第一次获得目标非常重要。
 2. 总体正确点击次数：围绕所有点击的总点击次数，最佳情况等于实际目标数量。我们的想法是计算最终选定目标中的正确点击次数。
- 不正确
 1. 首次不正确的点击：每个目标的第一次点击不正确，这意味着首次点击（目标数量）减去第一次正确点击。想法是知道第一次点击的丢失率。
 2. 总点击不正确点击次数：所有点击次数中的总点击次数错误，包括首次错误点击和其他错误点击。我们的想法是计算最终选定目标中缺失目标的数量。
- 最小点击距离（如文本输入中的 MSD）：将点击转换为目标所需的最小操作数。这个想法是文本输入度量标准最小字符串距离的变形。因为点击技术只包括两个操作，即移动和单击，所以转换似乎更容易，例如重新定位和重新选择错过的目标。
- 速度（每分钟选定的目标，如文本输入中的 WPM / CPM）：参与者单击每分钟的目标数量。

我们的想法是在一分钟内测量最终有效选定目标，而不考虑点击次数或移动次数。

- 效率（每个目标的操作，如文本输入中的 KSPC）：它是操作数量与目标数量的比率。我们的想法是计算选择任务的效率，直观地了解每个目标需要执行多少操作。

- 错误率：

1. 首先单击错误率：第一次单击不会导致选择目标目标。想法是计算第一次点击的丢失率。因为我们希望通常使用最少的时间和频率来获取目标，所以在第一次点击时知道丢失率非常重要。

2. 错误率：想法是计算最终中缺失目标总数与所选目标数之比。

3. 错误点击错误率：表示错误操作的总量与目标数量的比率。我们的想法是测量错误操作总数与目标数量的比率。

4. 总错误率：错误操作总数与所有操作数的比率。我们的想法是在操作过程中获得错误率。

- 时间：

1. 总时间：计算从出现的第一个目标到最后一个被点击目标的时间。

2. 反应时间：计算从出现的目标到开始移动的光标的时间。

3. 操作时间：计算从开始移动的光标到单击目标的时间。

- 距离：距离在两个目标的中心之间，距离由我们的点击任务设计决定。

眼动微交互技术设计

杨璐溪 201728060185

项目背景

微交互，即交互的开始与结束控制在 4 秒以内。微交互操作简单、容易执行，时间短返回快，不会干扰用户原先正在执行的主任务。在人机交互领域中，微交互不同于交互设计中 UI 的设计，而更加关注于对实际的交互方式的定义。当前市场上产品的交互方式大多以触摸交互、语音交互为主，眼动交互还处于研发阶段，并未普及。眼睛作为人类与外界交互最多的感官之一，当某些事件发生时，人们总想要第一时间用眼睛去查看。基于眼睛的生理特征，利用眼睛完成交互可以是非常理想的。除此之外，眼动交互还可以解决一些不方便同时使用两只手进行交互的情况，区别于语音交互在一些不方便说话的场景中的不便，眼动交互可以是很安静很优雅的。因此我们想要利用眼动来为人机交互设计一种快速的即时的微交互方式，例如消息提示、来电提示等。

我们希望这个设计可以广泛应用于各种不同的设备端。由于智能手表的屏幕很小，对于设计是一项挑战，如果设计能够在智能手表上很好的展现，那么它在更大的屏幕上也可以表现得很好。以及智能手表佩戴在人们的手腕上，拥有“无处不在”的特性。因此我们首先选择了针对智能手表的屏幕来进行设计。而这一领域目前在学术上并没有很多进展，我们要对比现有的几种交互设计搜集信息来提出自己的设计方案。

研究方法与使用设备

在项目开始阶段，我进行了大量相关文献的阅读与整理，同时分析整理现有的设计与方法。在这一阶段我们应用到了学习过的桌面调研与竞品分析的方法，以及关注到了他人实验中的用户体验，从中分析用户潜在需求。

我们会应用到心理学中的实验心理学方法来进行实验设计，例如如何设计对比实验，如何对被试进行 ABBA 实验法分组等。以及对于眼动交互的设计实验实施，我们会使用 Dikablis 眼动仪（如图 1）进行实验，通过 Unity 将设计原型实现在智能手表上，然后应用眼动仪的相关软件 D-Lab 进行数据的采集与整理。最后使用 R 语言对我们得到的数据进行分析。



图 1 Dikablis 眼动仪

设计方案

眼动交互是这个方向涵盖很广，眼动交互的方式与应用场景也有很多。眼动交互可以利用眼睛在某一位置的停留时间来进行选取，还可以结合眼动姿势进行目标物体的移动，也可以利用眼睛的动作（眨眼）来进行交互。但由于眼动的弥达斯效应（Midas effect），不自主的眼跳会对眼动监测系统造成影响，因此大多数的眼动交互需要精确的校准，需要让设备知道用户在什么位置，与屏幕的正交角度如何等，这样的要求就对现实中的实现造成了难度。因此我们选择利用对眼动平滑追踪（smooth pursuit）的检测来作为我们的眼动监测方式。眼动平滑追踪是指眼睛跟随某一物体或沿着某一路径进行移动，得到一条平滑的追踪线路。利用眼动平滑追踪，只需让监测系统知道用户的眼动路径是怎样的，而无需精确的校准。

结合现有的这些技术与设计，我们想到充分利用眼动交互可以快速且即时的特点，结合一些能够展现此特点的场景（例如短信提示、邮件提示、日程提醒等），来为微交互给出一种眼动微交互可能。通过与导师及实验室成员们的讨论，我的设计想法得到重新考虑与整理，最终得到了一个利用眼睛不同方向的直线运动来进行交互的方法，这一雏形很像是一个烟花，因此我们命名这个设计为 Firework。如图 2。

Name: **Firework**

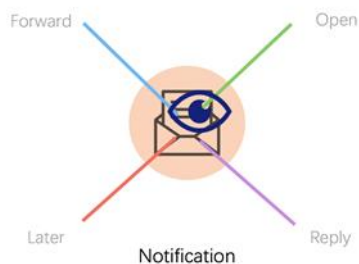


图 2 Firework 设计雏形

“这是一天的下午茶时间，你左手拿着一本小说，右手端着一杯暖暖的咖啡，缓缓送到嘴边，而正是这个时候你的左手感受到了一丝震动，是你的手表提示你有一个消息。这时候你并不着急放下手里的两样东西，而是轻轻地转动了你的手腕看向屏幕，原来是同事发送了一条业务消息，接着，firework 适时地出现了，你跟随着出现在屏幕上向左下角移动的烟花轨迹告诉你的手表，10 分钟后再次提醒你，因为你要回去工作了不想忘记你还要及时处理这个业务。然后你合上了小说，拿起了咖啡，径直走回到你工作的地方。”

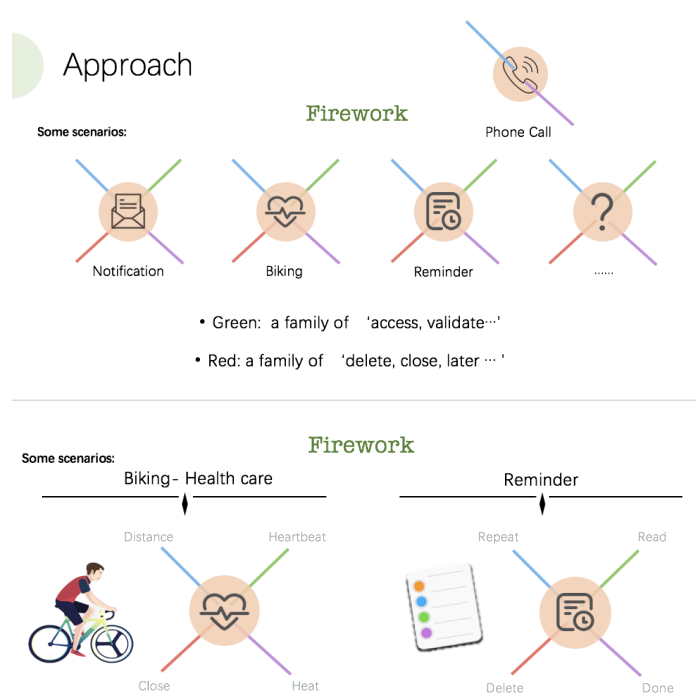


图 3 Firework 设计与可能的应用场景

我们希望这样的快速交互方式可以为人机交互带来更多可能性，因此我们也提出了一些可能的应用场景，如图 3。例如在骑行中，可以利用眼动快速查看心跳、距离、消耗能量等，也可以在使用提醒事项时快速浏览或者稍后提醒等。为了避免设计在每一个 APP 中都有不一样的含义，加大用户的记忆负担，我们为 Firework 定义了一些方向、颜色以及固定的含义，例如同一个方向的绿色都代表这打开、阅读等一类行为，而同一方向的红色都代表着关闭、删除等一类行为。

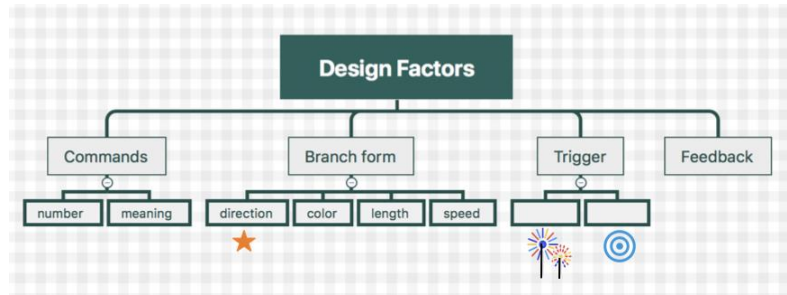


图 4 需要考虑的一些设计因素

在这个设计中我们考虑到了一些因素，如 Firework 可以完成的命令数目（方向数量），分支的形式，有无交互前的提示信息以及反馈。经过研究讨论，我们认为，对于设计中烟花能够在多少个角度上绽放可以放在实验中去测试，因此我们给了数量为 2、4、8 和 16 的方向和相应的区分度较大的颜色。对于有无提示，我们希望能够将其作为一个可平衡观测的量在实验中进行研究。

实验设计与实施

我们认为整个项目至少需要两个实验。首先，我们需要做的第一个实验是对这个设计进行评估实验。对于设计中的几个因素，我们需要分别进行考量。对于有无烟花的触发提示和眼动动作（平滑追踪/扫视）是我们决定进行平衡的两个设计因素。因此，我们的实验分为 4 个部分，共招募 12 个被试，每部分的实验时间约 40 分钟。

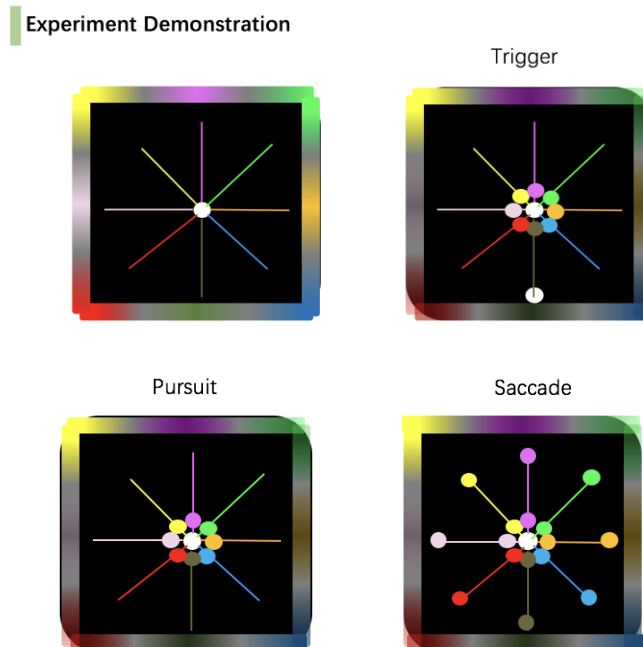


图 5 实验设计原型界面展示

实验的任务是要被试完成屏幕上出现的一句话中的指示（指示跟随某一颜色的小球运动）。小圆球运动的方式和速度均有三种模式，Trigger 的速度也有三种。我们需要通过实

验得到的数据对比这些设计因素，找到更好的结合模式。对于方向数量，我们也希望能够通过实验数据了解为这个设计我们最多可以给出多少方向（指令）。

我们的被试来自不同国家，实验开始前会帮助被试戴上眼动仪，并调整到合适的状态，以确保瞳孔运动可以被系统捕捉。每个实验时间大约 40 分钟，且眼动仪对于被试与屏幕的角度有一定要求，因此为了避免被试疲惫导致位置偏差影响实验，我们将手表放置在支架上，保持被试眼睛距手表约 40cm 进行实验。实验现场如下图。

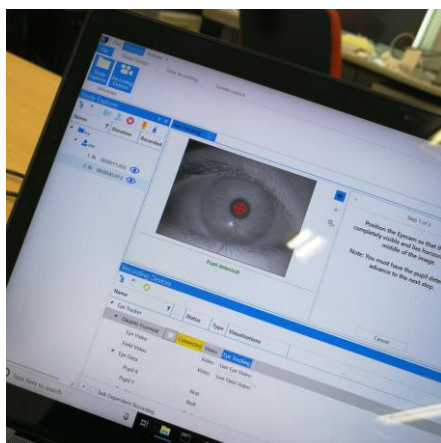


图 6 眼动仪调试



图 7 被试在进行实验

在实施实验的过程中，为了得到被试对于几个设计因素的反馈，以及对于整体任务的反馈。我们共设计了 3 个调查问卷。通过第一个问卷在实验开始之前得到一些被试的基本信息，实验结束后我们应用了 NASA_TLX 量表来得到他们对实验任务的相关评分，最后用一个较开放式的问卷来让被试说出他们的感受与偏好。

注：此项目隶属于 CHEC 实验室中眼动交互大方向下的一个子方向，意在为眼动交互提供更多可能。目前项目持续进行中，期待一个好的结果，也期待眼动交互方式能在不久的将来大展身手。