

北京師範大學

學士學位論文

論文題目：等比例變化圖形對一維視覺刺激比例估計的影響

院 系： 心理學部

專 業： 心理學

學 號： 201311061012

學 生 姓 名： 李昊宇

指 導 老 師： 蔣挺

指導教師职称： 講師

指導教師單位： 心理學院

2017 年 4 月

北京师范大学本科生毕业论文（设计）诚信承诺书

本人郑重声明： 所呈交的毕业论文（设计），是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

本人签名：

年 月 日

北京师范大学本科生毕业论文（设计）使用授权书

本人完全了解北京师范大学有关收集、保留和使用毕业论文（设计）的规定，即：本科生毕业论文（设计）工作的知识产权单位属北京师范大学。学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许毕业论文（设计）被查阅和借阅；学校可以公布毕业论文（设计）的全部或部分内容，可以采用影印、缩印或扫描等复制手段保存、汇编毕业论文（设计）。保密的毕业论文（设计）在解密后遵守此规定。

本论文（是、否）保密论文。

保密论文在_____年_____月解密后适用本授权书。

本人签名：

年 月 日

导师签字：

年 月 日

目 录

中文摘要	I
英文摘要	II
1 引言	1
1.1 比例估计的理论和误差	1
1.2 等比图形对估计的影响	1
1.3 实验假设	2
2 实验设计和方法	2
2.1 范式和因变量的测量	2
2.2 实验方法和流程	3
2.2.1 被试招募	3
2.2.2 数据处理	3
3 结果	4
4 讨论	4
参考文献	6

等比例变化图形对一维视觉刺激比例估计的影响

摘要 在图表制作过程中，使用长度或者面积的比例对数据进行形象表征非常常见，并且我们经常会使用到等比变化图形，所谓等比例变化图形指的是两个图形在直角坐标轴中所有点的坐标都以同一比例变化的图形，在这种图形中对一维视觉刺激的比例进行估计可能会受到图形的影响，而这个影响可以通过随机误差和系统误差加以体现。通过早期心理物理法实验的方法我们可以推论长度估计本身是不存在系统误差的。研究使用互联网游戏的形式进行基于距离估计的比例估计任务，一共收集了 6944 名被试的数据，结果表明，相比于仅仅呈现一维视觉刺激，等比例图形中一维视觉刺激长度的比例估计的随机误差更大；但是在系统误差当中，实验数据无法证明在两种情况下存在系统误差的差异，但是我们验证了在一般的比例估计中没有系统误差。实验结果表明我们应当避免在通过等比变化二维图形表示数据的时候使用一维视觉刺激作为指标。

关键词 比例估计 数据可视化 心理物理

The effect of scaled graph on the estimation of the ratio of two one-dimensional variables

Abstract When we make graphs, it is common to use length or space to demonstrate the ratio of one variable to another. When put into the scaled graph, the estimation of a one dimensional variable may be influenced. The coordinates of the scaled graph always change in the same scale when the graph shrinks or stretches. The influence can be resolved into systematic error and stochastic error. Adopting the method and theory of psychophysics, we can deduced that there is no systematic error in a plain ratio estimation. Using an internet based distance estimation game, the data of 6944 subjects are obtained. The result is that compared with only showing one-dimensional visual stimulation, estimation in scaled graphs are high in stochastic error. But as for systematic errors, none is found in both conditions. The results support that in scaled graphs, we should prevent from using one-dimensional visual stimulation such as length or angle to demonstrate our data.

Key Words Ratio estimation Data visualization Psychophysics

1 引言

在我们使用图表来表示数据的时候，我们实际上是把抽象的数值型变量变成了具体的长度，角度，或者是面积的比例。我们在进行这种转换的时候其实有一个假设，也就是我们对长度比例的估计是长度比例真实值的一个准确猜测，也就是说，我们的估计值和真实比例之间是不存在系统误差的，而我们的估计的不准确性是由于估计时候的随机误差所导致的。但是由于我们的估计存在随机误差，所以用无标识的图表的优势在于直观体现数据而不在于准确性，但是我们对这些比例的猜测的准确性保证了我们对数据直观体验的准确。

1.1 比例估计的理论和误差

认为系统误差不存在的观点只是我们通过经验做出的一个猜测，但是我们可以通过心理物理学的方法证明这一点。对于比例的估计实际上和物理量-心理量之间的转换是相似的。物理量的测量当中我们会使用一个固定的单位，例如在测量长度的时候我们使用米，而长度作为一个等比变量，我们使用它相对于固定的单位的比例作为物理量的测量。在心理量上，前人已经有研究证明长度的心理量和物理量之间的关系是线性的(Stevens & Galanter, 1957)。在这个研究之中，测量心理量的方法其实是使用数量估计的方式，也就是给定一个标准刺激让被试赋值，之后直接让被试给呈现的其他刺激使用数量赋值，这个方法本身其实使用的是被试的比例估计的能力，所以在长度估计的情况下，所证明的心理量和物理量的线性关系，实际上就是长度比例估计的心理量和物理量的相等。也就是说长度比例估计是不存在系统误差的。但是同样在心理物理的实验中，我们发现心理量的估计是十分不稳定的，所以长度比例估计的随机误差是必定存在的。

1.2 等比图形对估计的影响

我们在制作直方图的时候，常常使用面积来表示数据的多少，或者一些不常用的图表也会使用面积来表示数据。而一般使用的条形图，饼图等图表都是使用一维变量（长度，角度）来表示数据。当然饼图的情况比较特殊，其角度和面积的比例是一致的。在我们学习制作这些可视化图表的时候，规范告诉我们，使用某一些二维图形表示数据的时候，必须使用面积作为数据标识，而不能使用直径，边长这种一维指标。(Yau & 向怡宁, 2012)但是实际上二维变量和一维变量的比例不一致的时候，这两种变量的比例在人的感知当中会不会产生冲突呢？首先我们给这种一维变量和二维变量比例不一致的图形给出一个定义。我们考虑的二维图形的比例变化指的是在直角坐标系中两个坐标的等比例变化，也就是说面积的变化是单一坐标一维变量的变化比例的平方，我们把这种变化称为等长宽比变化。在对比两个等长宽比变化图形的时候，当我们不指导观察者特意去观察一维变量（例如长度）或是二维变量（例如面积）的变化的时候，观察者一般会报告二维变量的比例，这也是我们能使用面积去表示比例而不产生歧义的原因。但是当我们有意让被试报告等长

宽比变化的两个图形的一维变量的比例的时候，例如让被试报告两个圆的直径的比例，一维变量的比例会不会受到二维变量比例的干扰就是一个很有趣的问题。而这个问题主要关注的内容就是误差。

二维比例对一位比例的估计的影响主要有几个方面，一是一维变量出现在二维变量当中的时候，由于它没有被单独标识出来，那么他的显著性水平就会大大降低，必然就会导致估计过程中的随机误差的增大；其次，二维变量中的一维变量可能因为认知错觉的原因使得辨别的心理量和物理量之间产生系统误差，例如在正方形中长和宽的心理长度是不同的；再次，因为之前所说的在显示这一类等长宽比例变化图形的时候，面积估计是占主导的，那么虽然被试被指导去报告一维变量的比例，但是很难说二维变量的比例不会对这种报告产生影响。

1.3 实验假设

综合以上的理论，我们做出的理论假设是这样子的，首先由于在等长宽比变化图形中比较一维变量的时候，因为一维变量的显著性的降低，随机误差会上升；至于系统误差，我们通过心理物理的研究知道在单独呈现长度并要求估计的情况下，是不存在系统误差的，但是更为显著的二维变量可能会使得一维变量的比例估计结果变大。

2 实验设计和方法

为了验证等长宽比变化的图形中二维比例对一维比例估计的影响，我们简单地设置了一个自变量，也就是是否有干扰的二维变量，而因变量有两个，一个是估计的系统误差，一个是估计的随机误差。在无关变量的控制上，因为研究表明(王甦 & 方俐洛, 1964)，估计比例本身就会对估计比例的差别阈限产生影响，而差别阈限本身就是随机误差的一个影响因素，所以我们控制每一次估计的比例的大小是相同的。

2.1 范式和因变量的测量

在实验当中的比例估计采用了测距的比例估计范式。我们在观察物体的时候，物体的视角和距离的乘积等于物体的实际长度，那么同一个物体在不同的距离上的视角和距离是呈反比的。当视角使用弧度计算的时候，如果视角很小的话，可以证明弧度的大小近似地等于弦长，所以我们可以使用视野当中的长度来代替物体的视角进行计算。当给出一个物体 A 在视野中的长度和对应的距离，然后给出另一个相同物体在不同距离之下视野之中的长度，就可以根据比例估计得出 B 距离我们的距离。而通过收集被试距离估计的数据，就可以得到被试比例估计的结果。而如果将标准目标的比例设置为大于估计目标，然后将标准目标的距离设置为单位 1，然后估计目标的距离就是标准目标和估计目标视角弦长的比值，这样的设置相比于直接的数量估计任务并不会造成复杂的比例换算。

我们最终得到的被试估计数据和真实比例之间的差值可以分解成系统误差和随机误差

两部分，随机误差就是被试估计值和估计的平均值的离差的绝对值，而系统误差是被试估计的平均值和真实比例之间的差异。通过独立测量 t 检验可以检验在两个实验条件下随机误差的差异，而通过单样本 t 检验可以检验被试的估计数据和真实值之间是否存在系统误差。

2.2 实验方法和流程

实验方法采取将范式中所说的比例估计任务包装成一个射击游戏的形式，游戏当中的任务大体如下：

假设被试为一名坦克炮手，面前呈现一个瞄准镜，瞄准镜的中央标识被试调整方向机，炮弹在水平位置上上射击命中的点，在点下方，依次呈现调整被试调整高低机炮弹在竖直方向上对应距离的落点，水平轴上落点标识的距离和它距离 0 米刻度的在瞄准镜上的距离是成正比的，这一点和真实的瞄准镜是相同的，因为在现实当中炮弹在下落过程当中是可以忽略空气阻力的，所以炮弹下落的视角和射击正比的。被试通过瞄准镜可以看见模拟坦克在模拟环境当中行驶的截图，截图使用 War Thunder 游戏的坦克和环境的建模。通过手机运行游戏的时候在屏幕上看到的坦克的视角和实际上同型号坦克在标称距离上的视角相似，以免被试通过标称距离视角的印象做出影响实验结果的判断。当被试按下屏幕上一个点的时候，屏幕上会显示同型号坦克以同样角度面对你的时候 100M 距离上的视角大小，根据自变量的水平不同，屏幕上分别会显示坦克的二维轮廓来进行二维比例干扰下的比较，或者仅仅只使用直线表示坦克的视角长度，在仅仅显示直线的情况下，瞄准镜中的目标坦克上会有一条同样标识视角长度的直线。同样手机屏幕上观测的视角和实际坦克在 100M 处的视角相似。被试的任务是在一定时间之内，判断瞄准镜中坦克视角大小和 100M 标准大小的比例，从而判断远处的坦克的实际距离，然后拖动瞄准镜使用相应的距离标识瞄准坦克，开火之后屏幕上会显示是否击中。之后瞄准镜中的截图更换，被试进入下一试次。

以上描述的游戏使用 HTML5 Canvas 编程，并且上传到网络服务器当中，被试通过访问网址进行游戏，被试的游戏数据记录在服务器当中之后下载并进行分析。

2.2.1 被试招募

采用在社交平台传播游戏的方式进行招募被试，被试在进入游戏的时候填写昵称作为识别 ID，传播时以游戏得分作为排名，承诺最后排行榜前十的被试可以获得 10 元至 100 元不等的现金奖励，游戏计分和距离估计不同的一点在于，游戏计分也考虑到了水平方向上的射击误差，误差越小的被试得分越高。

游戏平台共开放两天，在结束当天晚上的 24 点公布排行榜并且发放奖励现金。

2.2.2 数据处理

采集到的数据在保存到本地之后使用 Matlab 制成表格之后使用 IBM SPSS Statistics 20

进行处理。对于每一名被试的数据，因为可以多次重复游戏，仅仅采取前 22 次试次的平均值作为该被试的估计值。系统误差的指标是被试估计距离和实际距离的差值的平均值，而随机误差的指标是被试估计距离和实际距离的差值和其均值之差的绝对值。

3 结果

实验共收集到 6944 名被试的结果，分别为干扰组 2999 名，非干扰组 3945 名，使用箱型图的外限作为标准排除极端被试，极端被试的来源可能是没有理解指导语内容或者随意游戏。余下干扰组 $N = 2607$ ，非干扰组 $N = 3367$ 。

表 1 两实验组的描述统计

组别		M	SD
干扰组	系统误差	1.26	19.01
	随机误差	12.33	14.47
非干扰组	系统误差	0.42	14.24
	随机误差	9.61	10.51

系统误差中，干扰组系统误差和 0 的 t 检验结果显著，但是效应量非常小($t = 3.39, p < 0.01, Cohen's d = 0.07$)，非干扰组系统误差和 0 的 t 检验结果不显著($t = 1.718, p = 0.086$)。随机误差中，干扰组和非干扰组的 t 检验结果显著($t = 8.44, p < 0.001, Cohen's d = 0.45$)效应量中等。

综合以上统计数据，我们可以看出是否有二维变量的干扰会增加对一维变量进行估计的时候的随机误差的大小，对于系统误差来说，在没有干扰的情况下，一维变量的估计是没有系统误差的，这验证了史蒂芬斯的心理物理学实验的结果。但是在有二维变量干扰的情况下，被试有可能倾向于低估比例的值，但是由于结果的显著性不高，而且效应量很小，导致不足以支持这一点，可能需要进一步的实验加以证明。

4 讨论

实验的结果和我们的预期部分相符，对于随机误差来说，我们预言二维变量的产生会使得需要对比的一维变量在物理上的刺激减弱，从而导致随机误差的增大。如果需要证明这一点，以后的研究可以通过控制在有二维变量的情况下，是否使得一维变量显著来检验。

而在系统误差方面，首先可以确定的是在仅存在一维变量比较的情况下，系统误差是符合理论预期不存在的，而存在二维变量干扰的情况下，目前的实验证据无法证明系统误差是否存在，而且实验中体现的趋势和我们的预测相反，在我们的预测当中，因为二维变量的比例是一维变量的平方（在实验中的标准刺激和变化刺激的比例都是大于 1 的），所以我们预期的可能的影响是二维变量的存在可能使得对与一维变量的估计夸大。导致这一点的可能有一下的两种情况：

其一是可能二维变量的存在可能并不会导致系统误差的产生，而我们实验中测量到的

系统误差可能是由于实验控制不严密所导致的，例如在实验当中，由于图片大小的原因对于每次瞄准的起始点的设置并没有特别控制，从而导致大部分的试次都是实际距离点在目标点的下方这样的设置就很有可能导致系统误差。

第二种可能是作为目标的图形，由于坦克的色彩有可能和背景的差别比较小，所以我们可能在知觉当中倾向于高估目标坦克的大小，而由于实验中的 100M 标准坦克的图形是通过线框勾画出来的，所以相比更为轮廓清晰。而所有的目标坦克的大小都是小于 100M 标准坦克的，所以导致了距离的低估。

在这个实验当中还有一些变量需要控制但限制于实验方法无法控制，有可能对实验结果产生影响，例如有研究()表明观测变量的高度有可能会对比例估计产生影响，被试进行观测时候的距离和姿势也有可能对比例估计产生影响(荆其誠, 彭瑞祥 & 方芸秋, 1963)，但是在本实验中，样本数量比较大，这些因素在两组之间产生差异的可能性非常低，所以额外加以控制的必要性也就没有那么高了。

总的来说，研究证明了在二维变量的呈现确实会对一维变量的比例估计产生影响，在数据可视化的书籍中所说的不要在等长宽变化的图形中使用一维变量来标明比例是有道理的，虽然这样做不一定会导致读者对于比例的估计产生偏差，但是有可能使得读者对于量的直观感觉的准确性下降，另外，如果图表中使用了一维变量在二维图形中表示数据，有没有醒目地标注这一点，读者的知觉可能让他错误地认为面积表明了变量的大小。

参考文献

- Bertamini, M., Yang, T. L., & Proffitt, D. R. (1998). Relative size perception at a distance is best at eye level. *Attention, Perception, & Psychophysics*, 60(4), 673-682.
- Fulton, S., Fulton, J., Ren, W., & Luo, Z. (2014). *HTML5 Canvas 开发详解*. 人民邮电出版社.
- Stevens, S. S., & Galanter, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *Journal of experimental psychology*, 54(6), 377.
- Yau, N., & Xiang, Y. (2012). *鲜活的数据: 数据可视化指南*. 人民邮电出版社.
- 荆其誠, 彭瑞祥, 方芸秋. (1963). 距离, 观察姿势对大小知觉的影响. *心理学报*, (1), 20-30.
- 郭秀艳. (2004). *实验心理学*. 人民教育出版社.
- 王甦, 方俐洛. (1964). 刺激的绝对大小对视觉长度比例辨别的影响. *心理科学*, (2), 55-61.