

数据 - 笔墨比率对过程监控任务绩效的影响^{*}

杨家忠^{**1,2,3} 张侃²

(¹中国民航飞行学院航空心理学教研室, 广汉, 618307) (²中国科学院心理研究所, 北京, 100101) (³中国科学院研究生院, 北京, 100039)

摘要 以 24 名参加飞行员心理选拔的大学生为被试, 采用模拟过程监控任务, 以反应时、心理负荷与情境意识为指标, 考察了数据 - 笔墨比率在动态图形显示设计中的应用。研究结果表明, 低数据 - 笔墨比率显示条件下的反应时较短、情境意识较高, 即适当增加表达数据点之间关系的笔墨虽然会降低数据 - 笔墨比率, 但有助于监控绩效。实验结果说明, 在将静态图形显示的设计原则运用于动态显示设计时, 应对其适用性进行检测。

关键词: 数据 - 笔墨比率 过程监控 心理负荷 情境意识

1 问题提出

随着计算机科学和技术的发展, 图形显示越见广泛用于印刷媒体与工业生产。对于印刷媒体如何利用表格、图示与图表等恰当地描述静态的定量数据, 研究者已提出许多设计原则^[1-3]。Cleveland 将表示数据的视觉格式塔称为图符 (graphic specifier), 图符描述数据的效果具有层次性, 因此在设计时应尽量选择处于高层次的图符来表达数据^[4]。Tufte 提出数据 - 笔墨比率 (data - ink ratio) 的概念, 即表达数据点的笔墨与所使用的总笔墨之比, 认为在设计时, 应遵循高数据 - 笔墨比率原则, 即尽可能删除不表达数据点的笔墨, 如过多的刻度标记和标签、冗余信息等^[5]。Gillan 与 Richman 的研究发现^[6], 尽管与作业任务类型、图示类型等存在交互作用, 总体而言, 数据 - 笔墨比率越高, 进行系列判断的时间越短, 准确性越高。这些结果使得对数据 - 笔墨与工作绩效的关系成为后来国际上的一个热点。Carswell 在一项元分析研究中比较了图符层次与数据 - 笔墨比率原则的预测效力, 发现图符的层次性能更好地预测作业绩效, 同时认为数据 - 笔墨比率最大化原则有点扩大化了^[7]。

工业生产中的图形显示主要用于描述系统的动态变化与状态, 适用于静态图形显示的设计原则在动态显示设计中的适用性如何较少受到关注。Blasio 在一项动态图形显示的研究中发现高数据 - 笔墨比率显示条件下的绩效较好, 支持高数据 - 笔墨比率原则^[8], 但实验中三类仪表的正常范围与异常范围值都是相同的。在低与中数据 - 笔墨比率显示条件下, 不必要的视觉成分 (如刻度标记、标签和颜色编码等) 降低了视觉搜索的速度, 从而影响了作业绩效。而实际上, 工业生产过程监控任务中所使用的

仪表通常具有不同的正常值与异常值范围。

在诸如航空、核电等复杂、动态变化的信息环境中, 操作者的心理负荷和情境意识 (situation awareness, 简称 SA) 与其作业绩效密切相关^[9]。心理负荷指单位时间内人所承受的心理加工的量, 涉及任务的心理资源需求; 情境意识涉及操作者在特定的时间和空间内对环境各种要素的知觉, 对其意义的理解, 及预测它们随后的状态^[10]。这两者都影响到操作者的作业表现, 前者是心理加工量的体现, 后者是心理加工的内容的反映。因此, 动态显示界面的设计和评价, 应考虑界面对操作者心理负荷与情境意识的影响, 良好的显示界面应兼具低心理负荷与高情境意识的特点。基于此, 众多的研究者都建议在系统评价时应同时测量心理负荷和情境意识^[11]。由于情境意识测量方法的敏感性不同, 研究者也建议, 如果可能的话应使用多种方法来测量情境意识^[11]。Endsley 提出的全面评估技术 (Situation Awareness Global Assessment Technique, 简称 SAGAT) 和 Taylor 提出的评定技术 (Situation Awareness Rating Technique, 简称 SART) 是测量情境意识的两种较为常用的方法^[12]。

本研究采用 Blasio 的研究范式^[8], 但实验中三种仪表的正常范围与异常范围不同, 更接近现实的工业条件, 增加任务的判断成分, 通过比较被试在不同水平数据 - 笔墨比率显示条件下的行为表现 (反应时为指标)、心理负荷与情境意识水平, 考察高数据 - 笔墨比率原则在动态显示设计中的适用性。

2 方法

2.1 被试

24 名参加飞行员心理选拔的男性大学生, 平均年龄 23 ($SD = 1.18$), 均为右利手, 无色盲。

* 此处的笔墨对应于印刷媒体使用的墨水或电子文本中的像素

** 通讯作者: 杨家忠, 男。E-mail: yangjz@psych.ac.cn

2.2 实验任务和程序

实验为模拟过程监控任务。在实际生产中,操作者必须监控系统状态变量,并在系统变量偏离预定值或正常范围时做出反应,或根据未来需要,对系统变量进行调整,这就需要操作者具有较高水平的情境意识,即保持对状态变量的意识。

工业生产中所使用的显示动态信息的仪表通常包括三个区域:正常区,警戒区和危险区。本实验中,以正常值范围标识正常区,异常值范围标识危险区。实验中每种显示条件下包括三组(压力、温度和流量)共九个仪表,仪表值每隔 2.5 秒更新。其中压力仪表的正常值范围是 2 - 4,异常值范围是 5 - 6;温度仪表的正常范围是 15 - 45,异常值范围是 60 - 75;流量仪表的正常范围是 80 - 160,异常值范围是 200 - 240。

实验包括三种类型的界面,根据每种界面中可删除的笔墨来确定其数据 - 笔墨比率的高低,可删除的量越多,则其比率越低。在低数据 - 笔墨比率显示条件下(见图 1),仪表上附刻度标记,带有完整的刻度标签。以绿色背景标识正常范围,红色背景标识异常值范围。以黑线的高度标识当前仪表值,同时带数字窗。在中数据 - 笔墨比率显示条件下(见图 2),仪表刻度标签除最低点值外,仅涉及正常与异常值范围数值,同时配以灰线标识这两个范围,以黑色短线的位置标识当前仪表值,无背景色和数字窗。在高数据 - 笔墨比率显示条件下(见图 3),以文本形式说明仪表的正常与异常值范围,当前仪表值仅以数字窗的形式出现。(注:因排版需要,图片大小已经过调整)

被试的任务是监控各仪表,要求在发现仪表读

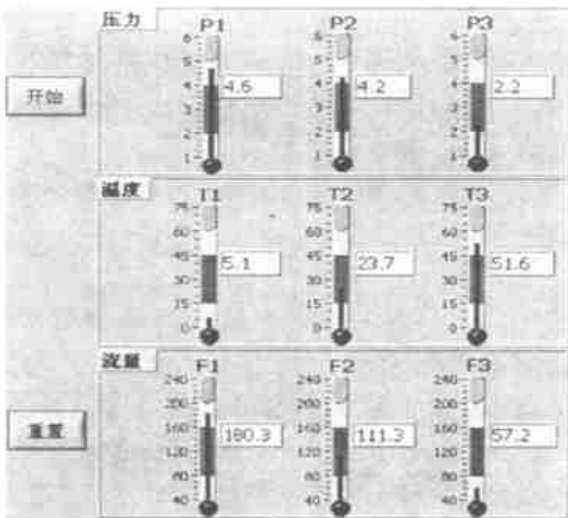


图 1 低数据 - 笔墨比率显示

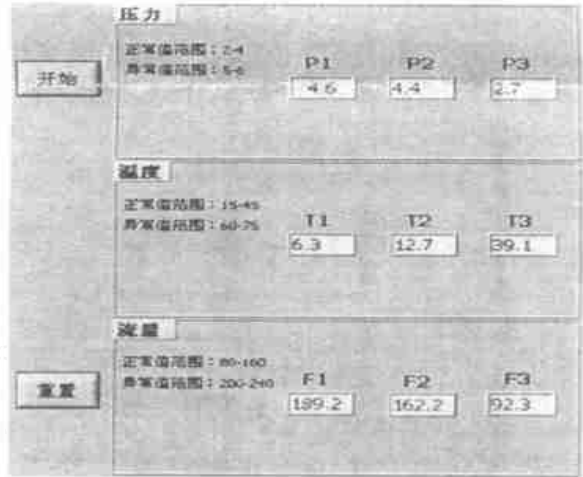


图 2 中数据 - 笔墨比率显示

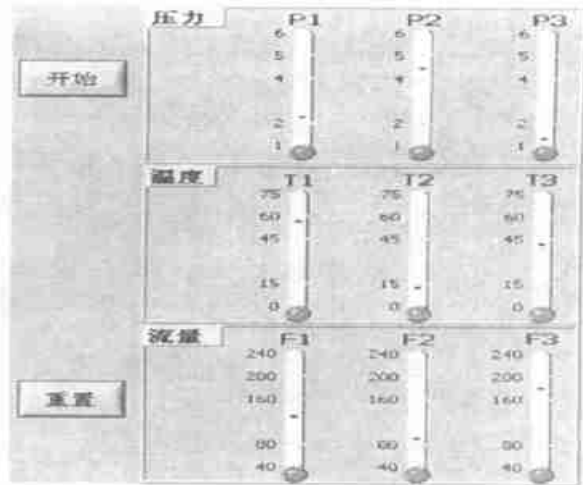


图 3 高数据 - 笔墨比率显示

数处于异常值范围时(称为“异常”事件),尽快按“重置”按钮,重新设定仪表。除了要对异常事件进行反应外,实验过程中,模拟任务会在随机选定的时间点停止,出现问题画面(称为“问题”事件),要求被试回答一个问题,以测量被试的情境意识。问题分定量和定性两种类型,定量问题的三个答案选项分别为正常值之下的区间范围、正常值范围,及正常值范围之上与异常值范围之下的区间范围;定性问题的答案选项是读数在正常范围之上、之中、还是之下。

根据 Endsley 对情境意识水平的三层次划分,定量问题测量水平 1 的情境意识,即对系统信息的觉察;定性问题测量水平 2 的情境意识,即对信息意义的理解。通过分析被试对定量和定性问题的回答情况可以确定其情境意识水平的高低,该方法为 SAGAT 测量。

每一测试阶段含 6 个随机选定的异常事件和 6 个随机选定的问题事件。12 个事件的时间间隔随机选自 27.5, 30, 32.5, 35 及 37.5 秒。测试过程中,12 个事件的出现顺序随机,每两个问题事件的平均间隔为 59.8 秒,接近 Endsley 提出的情境意识问题最小

时间间隔为 60 秒的建议^[10]。

出现异常事件后,被试按下“重置”按钮该事件才会消失。异常事件反应时的记录始于事件出现的时间点,止于被试按下“重置”按钮的时间点。

每种数据 - 笔墨比率显示条件下的异常事件与问题事件都是相同的,不同之处仅在于仪表的外观及事件的出现顺序存在差异。

实验前先向被试讲解并演示如何完成任务。实验包括三个测试阶段,每一阶段持续约 8 分钟。测试阶段之间被试可以自己安排作适当休息。在完成模拟监控任务后,让被试填写 NASA - TLX 心理负荷评定量表与 Taylor 的三维情境意识评价量表^[12]。心理负荷与情境意识皆采用 100mm 线法评定。

2.3 实验设计

实验采用 3(数据 - 笔墨比率:低、中、高) × 3(测试阶段:1、2、3)混合设计,其中数据 - 笔墨比率为被试间变量,测试阶段为被试内变量。每一数据 - 笔墨比率显示条件下随机分配 8 名被试。因变量为反应时、回答定量与定性问题的准确率、心理负荷及情境意识。

2.4 实验设备

实验在计算机上进行,使用 1024 × 768 像素彩色显示,程序用 Labview 7 编写,被试用鼠标按键反应。

3 结果分析

3.1 反应时

三种类型数据 - 笔墨比率显示条件下监控任务的反应时见图 4。重复测量分析表明存在显著的数据 - 笔墨比率主效应 ($F(2, 21) = 18.06, p < 0.001$);测试阶段的主效应不显著 ($F(2, 42) = 2.98, p > 0.05$),提示完成该实验任务没有练习效应;数据 - 笔墨比率与测试阶段之间的交互作用不显著 ($F(4, 42) = 0.72, p > 0.5$)。不同数据 - 笔墨比率条件下反应时的多重比较发现,低数据 - 笔墨比率与

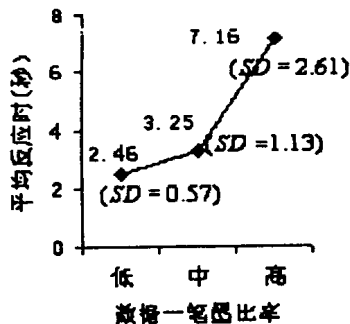


图 4 数据 - 笔墨比率对反应时的影响

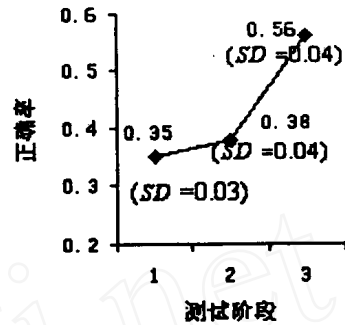


图 5 测试阶段对正确率的影响

中数据 - 笔墨比率组的反应时显著快于高数据 - 笔墨比率组,而前两者之间无显著差异。从反应时指标看,提示在动态显示设计中,低数据 - 笔墨优于高数据 - 笔墨比率。

3.2 正确率

本实验中,SAGAT法测量情境意识的问题分定量与定性两类,先比较两类问题对正确率的影响。One - way ANOVA 方差分析表明,问题类型对正确率的影响不显著 ($F(1, 430) = 0.78, p = 479$)。因此,下面正确率的分析对问题类型不作区分。

在正确率方面,重复测量分析表明数据 - 笔墨比率主效应不显著 ($F(2, 21) = 0.55, p > 0.5$);测试阶段的主效应显著 ($F(2, 42) = 12.65, p < 0.001$);两因素之间的交互作用不显著 ($F(4, 42) = 0.59, p > 0.5$)。不同测试阶段下的正确率见图 5,Fisher LSD 多重比较发现,测试阶段 3 的正确率显著大于测试阶段 1、2,提示存在学习效应。

3.3 主观心理负荷

采用 David 和 Laura 提到的方法^[13],根据被试在 NASA - TLX 6 个维度上的评分及每个维度的权重计算总的心理负荷。即将 6 个维度的评定值与其权重相乘,得到心理负荷的加权和,然后再除以 15,得到被试最终的心理负荷指数。方差分析表明,数据 - 笔墨比率的主效应不显著 ($F(2, 21) = 1.60, p > 0.1$),其中 M 低 = 68.88 ($SD = 11.43$), M 中 = 59.99 ($SD = 11.65$), M 高 = 66.38 ($SD = 7.00$)。

3.4 情境意识

Taylor 三维情境意识评价量表中,情境意识的计算采用公式^[12]:SA 得分 = 情境理解 - (注意资源的需求 - 注意资源的供给)。方差分析表明,3 种数据 - 笔墨比率条件下情境意识得分差异显著 ($F(2, 21) = 10.06, p < 0.001$)。进一步的多重比较发现,SA 低 ($M = 88.38, SD = 21.80$) > SA 中 ($M = 60.88, SD = 12.72$) > SA 高 ($M = 41.25, SD = 20.56$) (注“>”表示在 0.05 水平下差异显著),表明低数据 - 笔墨比率显示更有利于被试保持高水平的情境意识。

4 讨论

4.1 反应时

低与中数据 - 笔墨显示条件下,被试的反应远远快于高数据 - 笔墨比率显示。其原因在于低与中数据 - 笔墨比率的正常值与异常值范围都标识在仪表中,增加了适当的笔墨,使数据点之间的关系得到更好的描述,被试在完成任务时使用外界中的知识(knowledge in - the - world)^[14],只需要搜索标识仪表当前值的黑竖线(低数据 - 笔墨比率显示)与黑色短线(中数据 - 笔墨比率显示)是否处于异常值范围即可完成反应时任务。尤其是在低数据 - 笔墨比率显示条件,正常值范围与异常值范围采用突显的颜色编码。除了 Cleveland 提到的七种基本的图符外,研究者也常将客体显示器(object display)中运用的突显特征(emergent features)称为图符^[8],认为这些图符能够提高系统状态和变量值的可视化程度。

而在高数据 - 笔墨比率显示条件下,被试需要记住三种仪表的正常值与异常值范围,记忆负荷大,必须使用头脑中的知识(knowledge in - the - head),其作业除了搜索任务外,还必须结合头脑中的知识进行判断,心理操作的次数比另外两种条件多,不利于监控绩效^[14]。值得注意的是,低与中数据 - 笔墨比率显示除了使用相同的图符(即基准线上的线性显示)和数据 - 笔墨比率不同外,低数据 - 笔墨比率显示还使用了突显图符,但是其反应时差异不显著,这说明反应时的差异更多受图符的影响,这与前人的研究相似,即图符较数据 - 笔墨比率对绩效的影响更大^[7]。

4.2 心理负荷

单就异常事件判断任务讲,从反应时数据分析看,低与中数据 - 笔墨条件下的反应时任务要简单些,被试体验到的心理负荷似乎应该低些。但是三种显示条件下的问题任务对被试来讲过于困难,可能影响到被试对心理负荷的整体评价,而导致最终差异不显著。另外,尽管实验任务很困难,但被试在三种显示条件下的心理负荷得分并不是很高,其原因可能在于当作业负荷超过被试的心理限度后,被试会有意无意地降低绩效标准,致使负荷体验下降。

4.3 情境意识

实验采用 Endsley 提出的 SAGAT 技术,通过定量与定性问题来测量被试 1 与 2 水平的情境意识。被试的正确率在高、中、低三种显示条件下均无显著差异,表明 SAGAT 技术不能够敏感地反映被试情境意识水平的差异。由于需要在同时监控九个仪表的基础上,对某一仪表的值做出定量或定性的判断,被试

普遍认为该任务过于复杂和困难,三种显示条件下的正确率($M_{低} = 0.40, SD = 0.12; M_{中} = 0.45, SD = 0.11; M_{高} = 0.45, SD = 0.12$)只是稍高于机遇水平。SAGAT 基于记忆探查,要求操作者我报告记忆中的内容,由于数据的搜集方法与情境意识的大多数理论一致,几位研究者认为 SAGAT 测量技术构念效率高^[10]。但是有研究者特别提到工作记忆的局限问题^[9],认为在复杂的信息环境中,让被试进行记忆探查很容易受到这种限制,本实验的结果反映出了此种局限。

利用 SART 测量被试在三种显示条件下的情境意识则表明,在动态显示设计中,低数据 - 笔墨比率显示有利于情境意识的保持。SART 技术属于自我主观评定,一项关于情境意识测量的元分析研究表明^[11],SART 测量的敏感性较高。一些研究者认为情境意识是一个多维的结构,工作负荷是其不可分割的一部分,但是批评者指出在 SART 测量中包含心理负荷成分,可能会使 SA 与心理负荷混淆^[12]。已有研究表明 SART 测量中的情境意识得分与绩效测量相关,但是不清楚这到底是因为心理负荷成分还是情境理解的成分^[12]。本实验结果与该结论的前一点吻合,即被试在低与中数据 - 笔墨比率显示条件下的反应时较短,其情境意识水平也较高。另外,被试在三种显示条件下的心理负荷差异不显著,但其情境意识水平却存在显著差异。对情境意识量表中的情境理解维度进行方差表明低与中数据 - 笔墨比率显示有助于情境理解($F(2, 21) = 5.23, p < 0.05; M_{低-情境理解} = 73.25, M_{中-情境理解} = 61.50, M_{高-情境理解} = 53.88$),提示在 SART 测量中,情境意识与绩效的相关可归结于情境理解成分。

5 小结

5.1 本实验结果表明,复杂的过程监控任务条件下,低数据 - 笔墨比率显示设计较优,即适当增加表达数据点之间关系的笔墨有助于监控。

5.2 在将静态图形显示的设计原则运用于动态显示设计时,应对其适用性进行检测。

5.3 本实验证实 Taylor 提出的 SA 评定技术(SART)较 Endsley 提出的 SA 全面评估技术(SAGAT)更为敏感,提示 SART 测量中,情境意识与绩效的相关可归结于情境理解成分。

5.4 在本实验中,除了数据 - 笔墨比率不同外,还存在所使用图符的差异,其对绩效的重要性可能不同,提示在动态显示设计中,图符较数据 - 笔墨比率具有更高的预测效力。

6 参考文献

- 1 Carswell, C. M. Frankenberger, S. & Bernhard D. Graphing in depth: perspectives on the use of three dimensional graphs to represent lower - dimensional data. *Behaviour and Information Technology*, 1991, 10 (6) : 459 - 474
- 2 Carswell, C. M. & Wickens, C. D. Information integration and the object display: an interaction of task demands and display superiority. *Ergonomics*, 1987, 30 (3) : 511 - 527
- 3 Siegrist, M. The use or misuse of three - dimensional graphs to represent lower - dimensional data. *Behaviour and Information Technology*, 1996, 15 (2) : 96 - 100
- 4 Cleveland, W. S. & McGill, R. Graphical perception and graphical methods for analyzing scientific data. *Science*, 1985, 229: 828 - 833
- 5 Tufte, E. R. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press, 1983
- 6 Gillan, D. J. & Richman, E. H. Minimalism and the syntax of graphs. *Human Factors*, 1994, 36(4) : 619 - 644
- 7 Carswell, C. M. Choosing specifiers: an evaluation of the basic tasks model of graphical perception. *Human Factors*, 1992, 34 (5) : 535 - 554
- 8 Blasio, A. J. & Bisantz, A. M. A comparison of the effects of data - ink ratio on performance with dynamic displays in a monitoring task. *International Journal of Industrial Ergonomics*, 2002, 30 (1) : 89 - 101
- 9 Vidulich, M. A. Mental workload and situation awareness: essential concepts for aviation psychology practice. In: Tsang, P. & Vidulich, M. *Principles and practices of aviation psychology*. Mahwah, NJ: Erlbaum, 2003:115 - 146
- 10 Endsley, M. R. Measurement of situation awareness in dynamic systems. *Human Factors*, 1995, 37 (1) : 65 - 84
- 11 Vidulich, M. A. Testing the sensitivity of situation awareness metrics in interface evaluations. In: Tsang P, Vidulich M, *Principles and practices of aviation psychology*. Mahwah, NJ: Erlbaum, 2003:227 - 246
- 12 Endsley, M. R. & Garland, J. G. *Situation awareness analysis and measurement*. Mahwah, NJ: Erlbaum, 2000
- 13 David, B. & Laura, S. Age differences in perceived workload across a short vigil. *Ergonomics*, 2002, 45(13) : 949 - 960
- 14 Wickens, C. D. Gordon, S. E. & Liu, Y. L. *An introduction to human factors engineering*. New York: Addison - Wesley, 1998:223 - 232

The Effect of Data-ink Ratio on Performance in a Process-Monitoring Task

Yang Jiazhong^{1,2,3}, Zhang Kan²

(¹ Teaching and Research Section of Aviation Psychology, Civil Aviation Flight University, Guanghan, 618307)

(² Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101)

(³ Graduate School of the Chinese Academy of Sciences, Beijing, 100039)

Abstract Through a simulated process-monitoring task with reaction time, mental workload and situation awareness as indices, this study investigated the applicability of data-ink ratio in the design of graphs to the design of dynamic graphic displays. 24 college students attending the pilot selection were tested. The result indicated that, for the low data-ink ratio condition, the response time was much shorter, and the score for situation awareness was higher. With regard to the design of displays for complex process-monitoring tasks, increasing appropriately the amount of ink that can depict the relation between data points, despite the possibility of lowering the data-ink ratio, could be conducive to attaining better monitoring performance. The result indicates that a simple translation of the guidelines from a static to dynamic environment is not a guarantee for improved performance. It is necessary to test their applicability when they are applied to dynamic displays.

Key words: data-ink ratio, process monitoring, mental workload, situation awareness