

# 听觉注意负荷对视听整合的影响

毕浚皓, 李君缘, 平航, 罗瑞, 刘铸汉

贵州中医药大学人文与管理学院, 贵州 贵阳  
Email: 971425930@qq.com

收稿日期: 2021年7月2日; 录用日期: 2021年8月4日; 发布日期: 2021年8月13日

## 摘要

由于听觉负荷控制范式的不同, 听觉负荷对视听整合影响尚存争议。本研究采用RSVP范式操纵听觉负荷(低负荷、中负荷、高负荷), 通过设置0%、30%和70%三种干扰刺激同时呈现比例考察不同负荷对视听整合的影响。结果发现, 除视觉通道下高负荷反应时显著低于低负荷外, 其余通道不同负荷间反应时、击中率均无显著差异; 不同负荷间视听结合较单独视、听觉反应时更低、击中率更高; 竞争模型进一步研究表明, 低负荷下发生视听整合效应( $p < 0.05$ ), 中负荷下视听整合时程缩短、峰值降低、窗口期提前, 高负荷下未发生视听整合。结果表明, 视听整合效应受听觉负荷影响显著, 随负荷增加而降低直至消失。

## 关键词

听觉负荷, 视听整合, 竞争模型, 注意负荷

# Effects of Auditory Load on Audiovisual Integration

Junhao Bi, Junyuan Li, Hang Ping, Rui Luo, Zhuhan Liu

Medical Humanities College, Guizhou University of Traditional Chinese Medicine, Guiyang Guizhou  
Email: 971425930@qq.com

Received: Jul. 2<sup>nd</sup>, 2021; accepted: Aug. 4<sup>th</sup>, 2021; published: Aug. 13<sup>th</sup>, 2021

## Abstract

The effect of auditory load on audiovisual integration is still controversial due to the different auditory load control paradigms. To clarify the impact of auditory load on audiovisual integration, the RSVP paradigm was used to manipulate the auditory load in this study (low load, medium load, high load), and three kinds of interference stimulus presentation ratios of 0%, 30% and 70% were presented to investigate the effects of these on the audiovisual integration. The results showed

that there were no significant differences in response time and hit rate between different loads in visual and auditory channels; except that the response time under high load was significantly lower than that under low load in visual channel. The response time of audiovisual was lower than that of visual and auditory alone, but the hit rate, under different load. A further study of Race model shows that audiovisual integration effect occurs under low load ( $p < 0.05$ ), and the duration of audiovisual integration was shortened, the peak value was reduced, of course the window period was advanced under medium load. Under high load, no audiovisual integration occurred. The results showed that the audiovisual integration effect was significantly affected by the auditory load, and decreased until disappeared with the increase of the auditory load.

## Keywords

Auditory Load, Audiovisual Integration, Race Model, Attention Resources

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

感觉是个体存在的必要条件(Bexton, Heron, & Scott, 1954),但大量感觉信息若不能被高效加工,大脑将不堪重负。为提高效率,大脑将各种感觉信息进行整合。多感觉整合是指个体能够更有效地将来自不同感觉通道(如视觉、嗅觉、听觉、触觉等)的信息整合成统一高效的感知信号(McGurk & Macdonald, 1976; 顾吉有, 2016)。其中,视觉和听觉信息的整合被称为视听整合。例如双簧演出时,表演者引导个体将不同声音与外表整合以产生较好舞台效果。而注意对视听整合相当重要,研究表明注意资源的改变对视听整合效应有较大影响(Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010)。近年来,听觉作为影响注意的重要通道得到大量的研究和关注(Stoep, Janssen, Stigchel, & Nijboer, 2013)。

前人通常通过双任务范式(Dual tasks)研究听觉负荷对视听整合的影响(Fischler, 1976; Norman & Brown, 1975)。具体来看,Alsuis 研究中视觉刺激为女性读单词的视频片段,听觉刺激为单词音频,视听刺激为二者同时呈现;视觉负荷为叠加在视频上的无关图形,听觉负荷为叠加在听觉刺激上的无关音频。实验共分为有、无视听觉负荷两种情况,要求被试在无负荷条件下复述听到或看到的单词,负荷条件下除完成复述任务外还需探测重复出现的图片或音频(Alsuis, Navarra, & Soto-Faraco, 2007; Talsma et al., 2007; Tiippana, Andersen, & Sams, 2004)。结果表明负荷条件下,注意负荷减弱视听整合。顾吉有(2016)研究中视觉刺激为圆形和正方形图片,听觉刺激为“圆”和“方”音频,视听刺激为对应图片和音频同时呈现;听觉负荷为 1200 Hz 高音以及 400 Hz 低音。实验共分为有无负荷两种条件,其中无负荷条件下要求被试对圆形图片、音频及其同时呈现时按 F 键反应而对方形图片、音频及其同时呈现时按 J 键反应,负荷条件下除以上任务外还需对高音和低音分别按 D、K 键反应,以探究负荷条件下视听整合效应的变化(顾吉有, 2016)。结果表明视听整合较少受到听觉负荷的影响。

不论双任务范式下以何种方式设置干扰,其本质都是对注意资源中反应选择与执行阶段进行干预(吴彦文, 游旭群, 李海霞, 2014)。在这种方式下,对注意资源的控制通常需要改变任务内容,但任务内容的改变导致其对加工水平差异解释缺乏效力。快速系列视觉呈现范式(rapid serial visual presentation, RSVP)最早被 Broadbent 等人运用在注意瞬脱研究中(Broadbent & Broadbent, 1987),近年来又多被用于注意负荷研究(Talsma et al., 2007),即通过快速呈现干扰刺激以操纵注意资源。相对于双任务范式,该范式下通常

无需改变任务内容，同时研究设计更为简洁。

在 RSVP 范式中，研究者们常通过改变负荷刺激数量以及利用刺激间相似程度操纵注意负荷。具体来看，Fairnie 等人(2016)将刺激同时呈现在以被试为圆心的一个假想半圆中距离相等的五个点。其中目标刺激为狮子与狗的叫声，干扰刺激为鸭子、鸡、牛、乌鸦和公鸡叫声，并将实验分为仅目标刺激、目标刺激与一种非目标刺激、目标刺激与三种非目标刺激和目标刺激与五种非目标刺激同时呈现四个条件(Fairnie, Moore, & Remington, 2016)。Dalton 等人研究中听觉刺激以快速信息流的形式呈现，靶刺激为女性朗读的字母 P、T 音频，非靶刺激为男性朗读的字母音频。同时低负荷下听觉干扰刺激由 x 组成，高负荷下由 A、C、H、G、J、K 组成。不同负荷下均要求被试对靶刺激反应。结果表明听觉负荷对目标刺激加工无显著影响(Murphy, Fraenkel, & Dalton, 2013)。以上研究通过改变刺激数量或刺激间相似性为依据设置负荷，虽保持任务内容一致，但其不同负荷间刺激内容不一致同样影响其研究效度。Dalton 等人(2017)研究表明提高同时呈现刺激数量或强度将会降低对目标刺激的注意水平(Brungart, 2001; Dalton et al., 2017; Kidd, Mason, Richards, Gallun, & Urlach, 2007)，这为注意资源控制提供一种全新的思路，即通过改变刺激呈现水平对注意资源加以操控而无需改变任务内容或刺激内容。据此，本研究假设若通过特定听觉刺激作为掩蔽刺激与目标刺激同时呈现，将会影响对目标刺激的注意水平。

综上所述，本研究采用 RSVP 范式通过控制目标和干扰刺激出现比例操纵听觉负荷(低听觉负荷、中听觉负荷、高听觉负荷)，考察其对单独视觉、单独听觉以及视听整合的影响。基于前人研究结果，本研究预期听觉负荷对视听整合影响显著，中、高负荷下视听整合比例下降、峰值降低。

## 2. 方法

### 2.1. 被试

共二十九名被试(男性 9 名，女性 20 名)参与实验，年龄范围在 18~36 岁，平均年龄  $21.7 \pm 2.7$  岁。所有参与者均为贵州中医药大学本科生或研究生，听力和视力均正常，均为右利手，未患精神疾病，无脑部损伤史。以往没有参加过类似实验，结束后获得合理酬劳。

### 2.2. 实验材料

实验程序的编写和数据采集均采用 E-prime 2.0 (标准心理学实验软件)，刺激呈现在 Samsung S19E200 型号的液晶显示器上，可视尺寸 19 英寸，分辨率为  $1280 \times 1024$ ，刷新率为 75 Hz。视觉刺激被随机呈现在距离被试 60 cm 且背景为白色的显示器四个角(位于中心左侧或右侧  $12^\circ$ ，中心点上方或下方  $5^\circ$ )；听觉刺激通过电脑两侧扬声器呈现。

如图 1 所示，视听辨别任务中刺激共有 6 种呈现形式，其中视觉靶刺激(VT)为白色方格内有两个黑点的黑白棋盘图像( $52 \times 52$  mm，视角为  $5^\circ$ )，听觉靶刺激(AT)为 60 dB 字母 B，视听靶刺激(VAT)为二者的结合；视觉非靶刺激(VS)为黑白棋盘图像，听觉非靶刺激(AS)为 60 dB 字母 A, C, D, E, K, P, T 和白噪音。其中白噪音在低负荷不出现，中负荷出现次数为总试次的 30%的，高负荷出现次数为总试次的 70%。视听非靶刺激(VAS)为二者的结合。

### 2.3. 实验程序

实验采用 3 (低听觉负荷、中听觉负荷、高听觉负荷)  $\times$  3 (视觉通道、听觉通道、视听整合通道)的被试内实验设计。实验在被试完成 20 个试次的练习后正式开始。首先呈现一个黑色的固视点(持续 1000 ms)，呈现视、听刺激(持续 100 ms)，随后呈现一个固视点(持续 700~1000 ms)。低负荷条件下，对 VT 按左键进行反应，对 AT 按右键进行反应，VAT 出现时左右键一起按。对 VS、AS、VAS 不做任何反应。

中、高负荷下任务与低负荷一致，但白噪音与字母音频同时出现。实验要求被试在每个试次中，将视线集中于注视点，并对刺激做出既快又准的反应。实验共计 244 个试次，其中靶刺激与非靶刺激比例为 4:1。整个实验持续时间约 40 分钟。

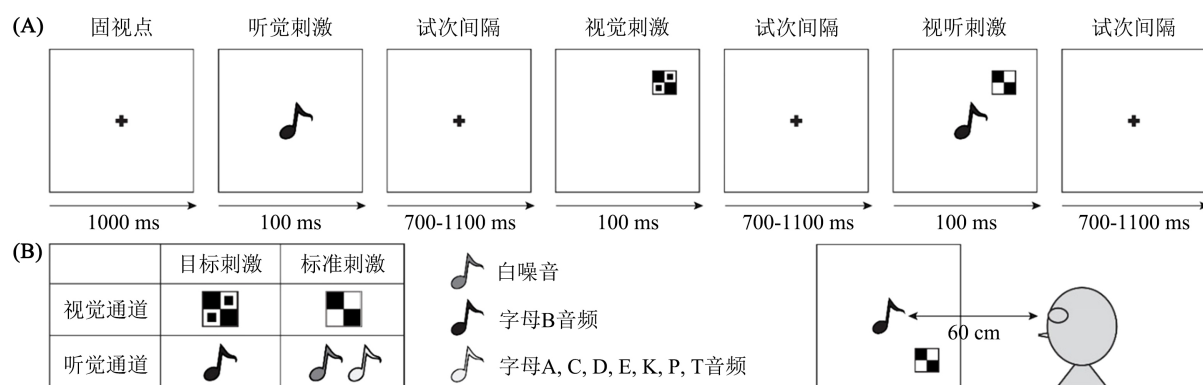


Figure 1. The paradigm of the experiment

图 1. 实验流程示意图

## 2.4. 数据分析

实验共 29 名被试，其中 9 人因一项或几项任务击中率低于或等于 0.75 而被删去。击中率是目标刺激正确次数与目标刺激总个数的百分比(Xu, Yang, Zhou, & Ren, 2020)。在低、中、高三个负荷条件下分别计算被试击中率与 RT。随后对击中率与 RT 进行 3 (低听觉负荷、中听觉负荷、高听觉负荷) × 3 (视觉通道、听觉通道、视听整合通道) 的重复测量方差分析，以检验负荷的有效性。

此外，为明确视听整合效应，使用累计分布函数(CDFs)分析 RT 数据并与竞争模型(Race model)比较。竞争模型是基于对单独视觉和听觉刺激的视觉和听觉反应的统计预测模型(Kahneman, 1973; Miller, 1986)。通过对该模型中视听整合模型概率与基于视觉、听觉预测概率的综合模型进行比较，若视听整合模型显著快于预测概率模型，则发生视听整合。最后，对竞争模型中视听整合模型与预测概率模型以 10 ms 为单位进行配对样本 T 检验，以明确各负荷下视听整合发生的时间。

## 3. 结果

### 3.1. 击中率

3 (低听觉负荷、中听觉负荷、高听觉负荷) × 3 (视觉通道、听觉通道、视听整合通道) 的重复测量方差分析表明负荷间主效应不显著  $F(2, 38) = 2.427, p > 0.05, \eta_p^2 = 0.113$  (见表 1)。感觉通道间主效应显著  $F(2, 38) = 29.492, p < 0.001, \eta_p^2 = 0.608$ ，进一步分析表明，视听结合击中率显著高于单独听觉和单独视觉( $p < 0.05$ )。负荷与刺激通道间交互作用不显著  $F(4, 76) = 0.511, p > 0.05, \eta_p^2 = 0.026$ 。

### 3.2. 反应时

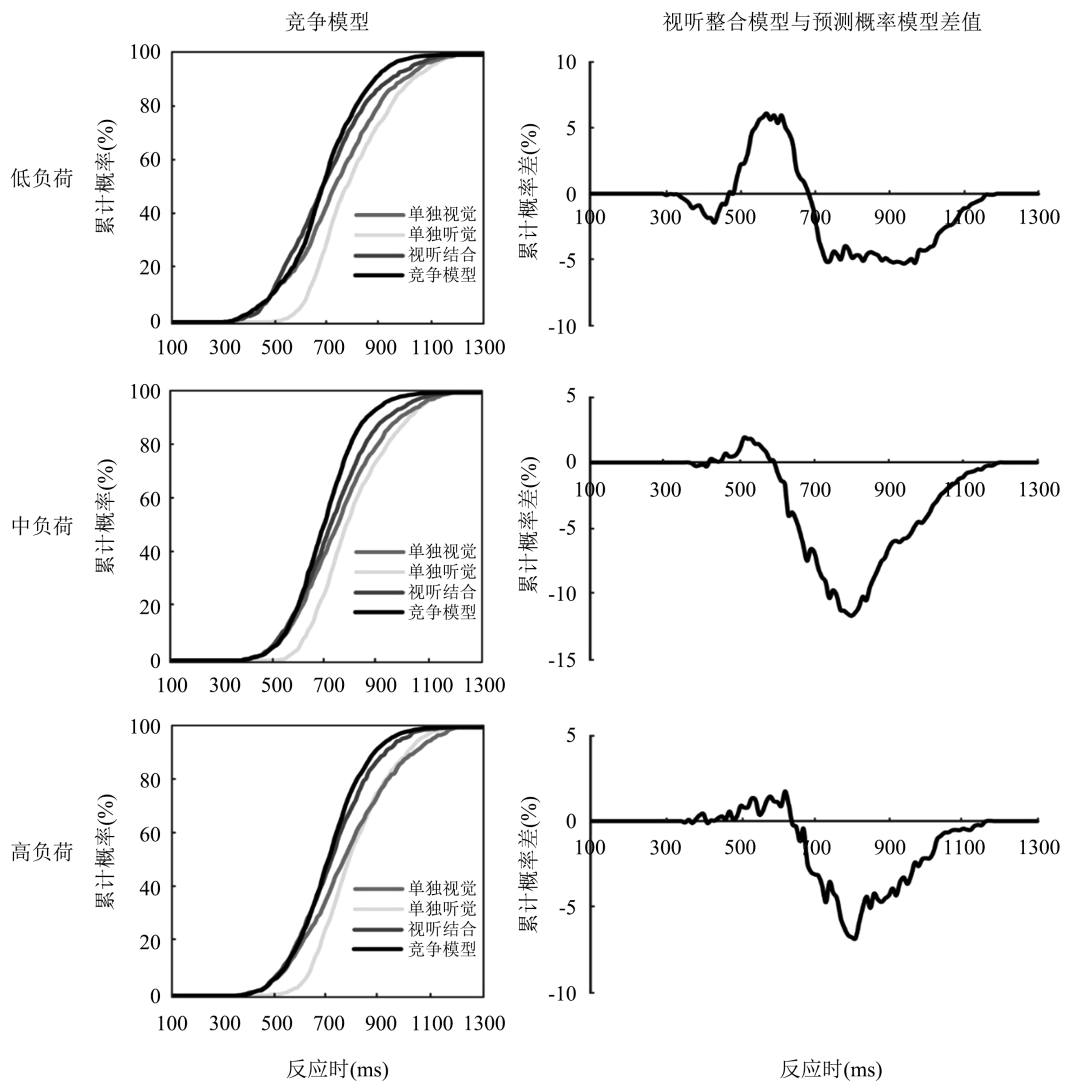
3 (低听觉负荷、中听觉负荷、高视觉负荷) × 3 (视觉通道、听觉通道、视听整合通道) 的重复测量方差分析表明负荷间主效应不显著  $F(2, 38) = 2.015, p > 0.05, \eta_p^2 = 0.096$  (见表 1)。感觉通道间主效应显著  $F(2, 38) = 16.997, p < 0.001, \eta_p^2 = 0.472$ ，进一步分析表明，视听结合反应时快于单独视觉和单独听觉( $p < 0.05$ )。负荷与刺激通道间交互作用显著  $F(4, 76) = 5.207, p < 0.05, \eta_p^2 = 0.215$ 。视觉通道下，低负荷反应时显著低于高负荷( $p < 0.05$ )。低负荷下，视听整合反应时显著低于单独视觉反应时( $p < 0.05$ )，单独视

觉反应时又显著低于单独听觉反应时( $p < 0.05$ ); 中、高负荷下, 视听整合反应时显著低于单独视觉、听觉反应时( $p < 0.05$ )。

**Table 1.** Response time (RT/ms), hit rate (%) and its standard deviation (SD) under different load conditions in audio-visual discrimination task

**表 1.** 视听辨别任务中不同负荷条件下反应时(RT/ms), 击中率(Hit rate/%)及其标准差(SD)

	低负荷			中负荷			高负荷		
	听觉	视觉	视听觉	听觉	视觉	视听觉	听觉	视觉	视听觉
RT	804.51	733.12	695.46	804.13	756.01	728.53	802.5	773.53	721.6
SD	94.71	131.62	116.44	65.57	111.36	98.25	62.68	126.54	106.23
Hit rate	93	94	100	93	94	99	92	92	100
SD	0.05	0.04	0.01	0.06	0.05	0.01	0.07	0.06	0.02



**Figure 2.** Competitive model and difference curve under different loads

**图 2.** 不同负荷下的竞争模型与差异曲线

### 3.3. 竞争模型(Race Model)分析

本研究利用不同通道间反应速度的差异采用 CDFs 对 RT 数据进行分析,并将视听结合 CDFs 与预测概率 CDFs 进行配对样本 T 检验。Race model 分析结果表明(见图 2),低负荷下视听整合效应显著的时间窗口为 100 ms (530~630 ms), $p < 0.05$ ;中负荷下视听整合效应显著的时间窗口为 10 ms (510 ms), $p < 0.05$ ;高负荷下未发生视听整合效应。同时,低负荷下视听整合效应峰值在 550 ms 出现,为 5.563%。

## 4. 讨论

本研究旨在明确简单视听辨别任务中听觉负荷对视听整合的影响,结果表明低、中负荷下发生视听整合。高负荷下未发生视听整合。同时中负荷下视听整合较低负荷时程缩短(10 ms),同时峰值降低(5.563%)、窗口提前。

研究表明视听整合效应随听觉负荷增大而降低直至消失,这与前人研究结果一致(Alsius, Navarra, Campbell, & Soto-Faraco, 2005)。在 Alsius 研究中视觉刺激为女性读单词的视频片段,听觉刺激为单词音频,视听刺激为二者同时呈现。同时要求被试对叠加在视频上的无关图形与叠加在听觉刺激上的无关音频做出反应。结果表明注意负荷减弱视听整合。可以预见,若将 Alsius 研究中视听觉负荷持续增大,视听整合效应水平将持续下降直至消失。这是因为大脑的协同作用(LaBerge, 1997)。具体来看,视、听觉刺激本应在枕颞叶区得到加工(Stein & Meredith, 1990; Teder-Slejrvi, McDonald, Russo, & Hillyard, 2002; 崔俭俭, 2020; 张迎, 2006),然而研究表明当注意指向一定活动时,可以改变相应的大脑功能区域或神经功能单元的激活水平,从而对当前认知活动产生影响。本研究中,由于听觉负荷水平的增加,导致大脑将注意资源更多用于负荷刺激加工而相对减少对目标刺激的注意,致使中、高负荷下视听整合效应显著降低直至消失。

与本研究假设一致的是中负荷下视听整合时程缩短、峰值降低,Alsius 等人研究同样表明听觉负荷下视听整合效应时程缩短、峰值降低。这是因为注意资源的限制(Kahneman, 1973),即中负荷下视听整合较低负荷所得到的注意资源更少,而为了保持对视听整合刺激的加工就不得不缩短对视听整合加工的时程并降低其峰值水平。

有趣的是,研究表明中负荷下视听整合窗口期提前,而以往研究并未对此要点进行深入探讨。具体来看,本研究低负荷较中负荷占用注意资源更少,对靶刺激加工能力更强,而加工窗口期却晚于中负荷。这是由于不同负荷下的注意机制的改变(Lavie & Tsai, 1994),Lavie 研究表明靶刺激的加工负荷决定干扰刺激被加工的程度:加工负荷高,靶刺激用尽了注意资源,干扰刺激不被注意,则出现早选择,反之出现晚选择。本研究中,中负荷下对注意资源的占用水平较低负荷更高,此时干扰刺激不被加工,发生早期选择。同时将更多注意资源集中于靶刺激,加快其反应,从而导致视听整合窗口期提前。

本实验完善了前人对于听觉负荷下视听整合的研究,发现听觉负荷对视听整合效应仍有较强影响;同时,该研究中听觉负荷的设置可能难度过大,未来可以进一步优化。

## 5. 结论

在以 RSVP 范式操控注意资源研究中,听觉负荷对注意资源有较大影响。视听整合效应随听觉负荷增大而减小直至消失。

## 参考文献

- 崔俭俭(2020). *不同难度的视觉任务认知决策过程脑机制的研究*. 硕士学位论文, 郑州: 郑州大学.  
顾吉有(2016). *注意对视听整合加工的影响*. 博士学位论文, 天津: 天津师范大学.

- 吴彦文, 游旭群, 李海霞(2014). 注意力资源限制与双任务的相互干扰机制. *心理学报*, 46(2), 174-184.
- 张迎(2006). *大脑视觉通路在不同注意任务难度下的ERP研究*. 硕士学位论文, 天津: 河北工业大学.
- Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to Touch Weakens Audiovisual Speech Integration. *Experimental Brain Research*, 183, 399-404. <https://doi.org/10.1007/s00221-007-1110-1>
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual Integration of Speech Falters under High Attention Demands. *Current Biology CB*, 15, 839-843. <https://doi.org/10.1016/j.cub.2005.03.046>
- Bexton, W. H., Heron, W., & Scott, T. H. (1954). Effects of Decreased Variation in the Sensory Environment. *The Canadian Journal of Psychiatry*, 8, 70-76. <https://doi.org/10.1037/h0083596>
- Broadbent, D. E., & Broadbent, M. (1987). From Detection to Identification: Response to Multiple Targets in Rapid Serial Visual Presentation. *Perception & Psychophysics*, 42, 105-113. <https://doi.org/10.3758/BF03210498>
- Brungart, D. S. (2001). Informational and Energetic Masking Effects in the Perception of Multiple Simultaneous Talkers. *Journal of the Acoustical Society of America*, 109, 2527-2538. <https://doi.org/10.1121/1.1408946>
- Dalton, P., Murphy, S., & Spence, C. (2017). Auditory Perceptual Load: A Review. *Hearing Research: An International Journal*, 352, 40-48.
- Fairnie, J., Moore, B., & Remington, A. (2016). Missing a Trick: Auditory Load Modulates Conscious Awareness in Audition. *Journal of Experimental Psychology Human Perception & Performance*, 42, 930-938. <https://doi.org/10.1037/xhp0000204>
- Fischler, I. (1976). Beating the Memory Drum: Rapid Sequential Visual Presentation (RSVP) on a 4K PDP-8. *Behavior Research Methods & Instrumentation*, 8, 233-234. <https://doi.org/10.3758/BF03201782>
- Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., & D Urlach, N. I. (2007). *Informational Masking*. New York: Springer.
- Laberge, D. (1997). Attention, Awareness, and the Triangular Circuit. *Consciousness & Cognition*, 6, 149. <https://doi.org/10.1006/ccog.1997.0305>
- Lavie, N., & Tsal, Y. (1994). Perceptual Load as a Major Determinant of the Locus of Selection in Visual Attention. *Perception & Psychophysics*, 56, 183-197. <https://doi.org/10.3758/BF03213897>
- McGurk, H., & Macdonald, J. (1976). Hearing Lips and Seeing Voices. *Nature*, 264, 746-748. <https://doi.org/10.1038/264746a0>
- Miller, J. (1986). Timecourse of Coactivation in Bimodal Divided Attention. *Perception & Psychophysics*, 40, 331-343. <https://doi.org/10.3758/BF03203025>
- Murphy, S., Fraenkel, N., & Dalton, P. (2013). Perceptual Load Does Not Modulate Auditory Distractor Processing. *Cognition*, 129, 345-355. <https://doi.org/10.1016/j.cognition.2013.07.014>
- Norman, D. A., & Bobrow, D. G. (1975). On Data-Limited and Resource-Limited Processes. *Cognitive Psychology*, 7, 44-64. [https://doi.org/10.1016/0010-0285\(75\)90004-3](https://doi.org/10.1016/0010-0285(75)90004-3)
- Stein, B. E., & Meredith, M. A. (1990). Multisensory Integration. *Annals of the New York Academy of Sciences*, 608, 51-70. <https://doi.org/10.1111/j.1749-6632.1990.tb48891.x>
- Stoep, N., Janssen, M. A., Stigchel, S., & Nijboer, T. (2013). Exogenous Auditory Cues Decrease Audiovisual Integration of Simple Lights and Sounds. *The 14th International Multisensory Research Forum 2013*, Jerusalem, 3-6 June 2013, 223-224.
- Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex*, 17, 679-690. <https://doi.org/10.1093/cercor/bhk016>
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The Multifaceted Interplay between Attention and Multisensory Integration. *Trends in Cognitive Sciences*, 14, 400-410. <https://doi.org/10.1016/j.tics.2010.06.008>
- Teder-Slejrvi, W. A., McDonald, J. J., Russo, F. D., & Hillyard, S. A. (2002). An Analysis of Audio-Visual Crossmodal Integration by Means of Event-Related Potential (ERP) Recording. *Cognitive Brain Research*, 14, 106-114. [https://doi.org/10.1016/S0926-6410\(02\)00065-4](https://doi.org/10.1016/S0926-6410(02)00065-4)
- Tiippana, K., Andersen, T. S., & Sams, M. (2004). Visual Attention Modulates Audiovisual Speech Perception. *European Journal of Cognitive Psychology*, 16, 457-472. <https://doi.org/10.1080/09541440340000268>
- Xu, Z., Yang, W., Zhou, Z., & Ren, Y. (2020). Cue-Target Onset Asynchrony Modulates Interaction between Exogenous Attention and Audiovisual Integration. *Cognitive Processing*, 21, 261-270. <https://doi.org/10.1007/s10339-020-00950-2>