

Dual-Task Performance Consequences of Imperfect Alerting Associated with a Cockpit Display of Traffic Information

C. Wickens., and A. Colcombe. (2007)

Human Factors, Vol. 49, No. 5, pp. 839-850.

Introduction

- 人間とシステムとのインタラクションが生じる状況
 - 人間は並列的な活動を行う (e.g., Dixon & Wickens, 2006)
- 著者の興味
 - 人間と航空事故を検出する警報システムとのインタラクション
- 人間に影響を与える要因
 - (1) 警報システムの信頼性
 - (2) 警報システムの閾値の設定
 - ◇ silent と alert の割合の設定
 - 高い hit rate → 高い false alert rate を招く
 - 高い silent rate → 高い miss rate を招く
- システムへの依存
 - 信頼性の高いシステムを使用する場合
 - ◇ 人間はシステムをフルに活用 (Parasuraman, 1987)
 - 信頼性の低いシステムを使用する場合
 - ◇ システムの信頼性低下に従って、人間はシステムを使用しなくなる
 - ◇ システムを全く使用しない方がよい場合でも、人間はシステムを使用し続ける (Maltz & Shinar, 2003; Wickens & Dixon, 2007)
- 本研究の興味
 - 警報システムのエラーの明確さによって、人間のシステムへの依存度は変化するか？
 - 警報システムの miss と false alarm (FA) は人間に異なる影響を与えるか？
- この研究で考慮する警報システムの 4 つの要因
 - (1) 警報システムの閾値の設定
 - ◇ miss と FA が与える影響
 - (2) 警報のモダリティ
 - ◇ 視覚警報と聴覚警報が与える影響
 - (3) 警報タイプ
 - ◇ 2 段階(yes-no)の警報と 3 段階(yes-maybe-no)の警報が与える影響
 - (4) 手動で行う課題の難易度

- 本研究で扱うシステム
 - 交通情報表示 (CDTI: cockpit display of traffic information)を使用
 - ◇ 航空機の操作画面に備え付けられた交通情報を表示するディスプレイ
 - 開発段階
 - ◇ パイロットは航空機を操縦し、警報システムは他機の接近を警報
 - 航空事故を防ぐ
- 仮説
 - 警報システムが FA 傾向の場合 (警報システムが miss 傾向の場合と比較して)
 - ◇ (1) 参加者の従順性が低下し(Meyer, 2001, 2004), 他機接近の検出が遅くなる
 - ◇ (2) 参加者の信頼性は増加し(Meyer, 2001, 2004), 手動で行う課題パフォーマンスが高くなる
 - 聴覚警報を使用する場合 (視覚警報を使用する場合と比較して)
 - ◇ (3) 他機接近の検出は向上する
 - (3a) 手動で行う課題のパフォーマンスは低くなる
 - Preemption (占拠, 無効) theory (Spence, 2001)
 - 聴覚情報に注意を奪われる
 - (3b) 手動で行う課題のパフォーマンスは高くなる
 - Multiple resource theory (e.g., Wickens, 2002)
 - 視覚情報と聴覚情報のそれぞれに注意を向ける
 - 3段階の警報を使用する場合 (2段階の傾向を使用する場合と比較して)
 - ◇ (4) 細分化された情報が表示されることから, 参加者は, 2つの課題間で適切な注意分配を行う

Experiment 1

Method

- パイロットの訓練生 20 名が実験に参加
- 図 1 は実験課題のディスプレイ
 - 二重課題
 - ◇ 手動で追従課題を行い, 警報システムに他機の監視を任せる
 - 追従課題 (中央)
 - ◇ 移動するターゲットの内側にカーソルを留める
 - CDTI (右上)
 - ◇ 警報システムが他機接近を検出した場合, 視覚的または聴覚的に警報される
 - ◇ 他機が接近している場合, CDTI 上の他機のアイコンをマウスクリック
- 表 1 は警報システムの閾値の設定

- 実験 1 では、miss 傾向の警報システムを使用 (FA と miss の比率は 1:1)
- 実験 2 では、FA 傾向の警報システムを使用 (FA と miss の比率は 4:1)
- 要因操作
 - 追従課題の難易度
 - ◇ stability(安定性) : stable(容易)/unstable(難しい)
 - ターゲットの安定性 (?)
 - 警報のモダリティー
 - ◇ 視覚/聴覚
 - 視覚警報 : 他機が接近した場合、画面中央の棒が赤に変色
 - 聴覚警報 : 他機が接近した場合、"conflict conflict" という機械音声
 - 警報タイプ
 - ◇ 2 段階(yes-no)/3 段階(yes-maybe-no)
 - 視覚警報 : maybe の場合、画面中央の棒が黄色に変色
 - 聴覚警報 : maybe の場合、"traffic traffic" という機械音声
- 手順
 - 全ての要因を参加者内実験で実施
 - 実験は 2 日に分けて行われた
 - ◇ 一方の日に警報タイプ 2 段階、もう一方の日に 3 段階で実験を実施
 - ◇ 各実験で、警報のモダリティー(視覚/聴覚)と課題の難易度(stable/unstable)の各 2 条件全 4 通りの組み合わせで課題を実施
 - 課題の練習(追従課題のみ、CDTI のみ、両課題)を実施し、その後本課題を実施
 - 本課題は約 14 分間
 - CDTI の設定
 - ◇ 画面には 4 機以上の飛行機が現れることはなかった
 - ◇ 10 秒ごとに新しい飛行機が画面に現れた

Results

- 分析された従属変数
 - 他機接近を検出した際の(マウスクリックの)反応時間
 - 他機接近の検出の敏感さ (d')
 - ◇ 他機の接近、未接近の識別の正確さ
 - 追従課題のパフォーマンスエラー
 - ◇ ターゲットとカーソルとの距離 (?)
- 各従属変数に関して、3 要因の参加者内分散分析を実施
 - 警報のモダリティー(視覚/聴覚)、警報タイプ(2/3 段階)、課題の難易度(stable/unstable)
- 検出の反応時間

- 課題の難易度要因に主効果 ($F(1, 11)=7.02, p<.05$)
 - ◇ **stable < unstable**
- 警報タイプ要因と課題の難易度要因に交互作用 ($F(1, 11)=8.62, p<.05$) (図 2)
 - ◇ 課題が **stable** な場合, 3 段階の警報では反応が速い
 - ◇ 課題が **unstable** な場合, 3 段階の警報では反応が遅い
- 検出の敏感さ (d')
 - 警報のモダリティ要因に主効果 ($F(1, 11)=54.07, p<.01$)
 - ◇ 視覚 < 聴覚
- 追従課題のパフォーマンスエラー
 - 課題の難易度要因に主効果 ($F(1, 11)=372.78, p<.01$)
 - ◇ **stable < unstable**
 - 警報タイプ要因に主効果 ($F(1, 11)=8.26, p<.05$)
 - ◇ 2 段階 < 3 段階
 - 警報のモダリティ要因と警報タイプ要因に交互作用 ($F(1, 11)=6.13, p<.05$) (図 3)
 - ◇ 視覚的モダリティを使用した 3 段階の警報ではエラーが多い

Experiment 2

Method

- FA と miss の比率を変化(実験 1-1:1, 実験 2-4:1)させて, 実験 1 と同様の実験を実施
- パイロットの訓練生 12 名が実験に参加

Results

- 実験 1 と同様の従属変数に関して同様の分析を実施
- 検出の反応時間
 - 警報のモダリティ要因に主効果 ($F(1, 11)=4.09, p=.07$)
 - ◇ 視覚 > 聴覚
 - 3 要因間に 2 次の交互作用 ($F(1, 11)=5.2, p<.05$) (図 4)
 - ◇ 課題が **unstable** 場合, 3 段階の視覚警報では反応が遅い
- 検出の敏感さ (d')
 - 警報のモダリティ要因に主効果 ($F(1, 11)=3.41, p=.06$)
 - ◇ 視覚 < 聴覚
 - 警報タイプと課題の難易度要因に交互作用 ($F(1, 11)=4.69, p<.05$)
 - ◇ 課題が **unstable** な場合, 3 段階の警報では反応が鈍い
(他機の接近, 未接近の識別が正確ではない)
- 追従課題のパフォーマンスエラー

- 課題の難易度要因に主効果 ($F(1, 11)=675.66, p<.01$)
 - ◇ stable < unstable
- 警報のモダリティ要因に主効果 ($F(1, 11)=3.59, p<.10$)
 - ◇ 視覚 < 聴覚

Between-Experiment Comparison

- 各従属変数に関して、警報システムの閾値要因(miss 傾向/FA 傾向)を加えた分析
 - 実験 1(miss 傾向)と 2(FA 傾向)の実験間の比較
 - 各従属変数に関して、4 要因の分散分析を実施 (?)
- 検出の反応時間
 - 警報システムの閾値要因と課題の難易度要因に交互作用 ($F(2, 22)=4.37, p<.05$) (図 5)
 - ◇ 課題が stable な場合、実験 1(miss 傾向)よりも実験 2(FA 傾向)で反応が遅い
 - 警報システムの閾値要因と警報タイプ要因に交互作用 ($F(2, 22)=3.42, p=.07$)
 - ◇ 3 段階の警報の場合、実験 1(miss 傾向)よりも実験 2(FA 傾向)で反応が遅い
- 検出の敏感さ (d')
 - 警報システムの閾値要因に主効果 ($F(2, 22)=65.71, p<.01$)
 - ◇ 実験 1(miss 傾向) < 実験 2(FA 傾向)
- 追従課題のパフォーマンスエラー
 - 警報システムの閾値要因に主効果 ($F(2, 22)=3.24, p=.05$)
 - ◇ 実験 1(miss 傾向) > 実験 2(FA 傾向)
 - 警報システムの閾値要因と警報のモダリティ要因に交互作用 ($F(2, 22)=4.64, p<.05$) (図 6)
 - ◇ 実験 2(FA 傾向)のときに追従課題のパフォーマンスは低下
 - ◇ 視覚警報のときにその低下の程度が大きい

Discussion

- 4 つの仮説の検証結果
- 警報システムが FA 傾向の場合 (警報システムが miss 傾向の場合と比較して)
 - 他機接近の検出が遅くなる (仮説 1)
 - ◇ 課題が容易な場合のみ仮説 1 を支持 (図 5)
 - 従順性が低下 (cry-wolf effect)
 - 検出の敏感さの分析から、参加者は敏感に反応 (他機の接近, 未接近を適切に識別)
 - ◇ 参加者は、適切に警報システムを活用
 - 手動で行う課題パフォーマンスが高くなる (仮説 2)
 - ◇ 仮説 2 を支持 (図 6)

- 信頼性が増加
- 聴覚警報を使用する場合（視覚警報を使用する場合と比較して）
 - 他機接近の検出は向上する（仮説 3）
 - ◇ 実験 2 の反応時間
 - ◇ 実験 1 と 2 の検出の敏感さの分析から
 - 仮説 3 を支持
 - 手動で行う課題のパフォーマンスは低くなる（仮説 3a）
 - ◇ 実験 2 の追従課題のパフォーマンスエラーの分析から
 - 仮説 3a(Preemption (占拠, 無効) theory)を支持
 - 様々な状況での検討が必要
- 3 段階の警報を使用する場合（2 段階の傾向を使用する場合と比較して）
 - 参加者は、2 つの課題間で適切な注意分配を行う（仮説 4）
 - ◇ 実験 2 の反応時間と検出の敏感さ
 - ◇ 実験 1 の追従課題のパフォーマンスエラーの分析から
 - 仮説 4 は支持されなかった

Conclusion

- 実験の結果から確認された内容
 - 警報システムの閾値が変化すると
 - ◇ システムへの従順性に影響して反応時間が変化
 - Miss 減少, FA 増加 → 従順性低下, 反応時間が遅くなる
 - ◇ システムへの信頼性に影響して手動で行う課題パフォーマンスが変化
 - Miss 増加, FA 減少 → 信頼性低下, 手動で行う課題パフォーマンス低下
 - 基本的に聴覚警報システムは有効
 - ◇ FA が増加するとその有効性は低くなる
- 3 段階の警報システムの有効性はみられなかった
 - 今後, どのような状況で有効であるか検討する必要
- 今回の研究の限界
 - 実際の飛行機とは以下 3 点で異なる
 - ◇ 追従課題：実際の操作感に劣る
 - ◇ CDTI：実際のディスプレイより小さい
 - ◇ 他機の接近：実際の接近機会よりはるかに多い
 - 実際の飛行機に近いシミュレータを用いた実験を推奨

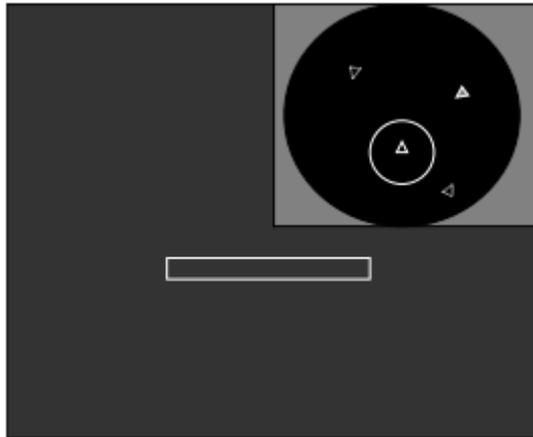


Figure 1. Display for the experimental task in Experiments 1 and 2 with the centrally presented tracking task and the CDTI in the upper right corner.

TABLE 1: Event Rates for the Two Experiments and Alert Conditions

Automation Response	Conflict	Nonconflict	$p(H)$	$p(FA)$
Experiment 1: High (Neutral) Threshold				
(a) Binary				
Yes	36	10	.78	.29
No	10	24		
(b) Likelihood				
Yes	30	7	Automation Yes = Hit: 30/47 = .64 (high-level)	Automation Yes = Hit: 7/36 = .19
Maybe	10	11	Automation Yes + Maybe = Hit: 40/47 = .85 (high + mid)	Automation Yes + Maybe = Hit: 18/36 = .50
No	7	18	Automation Yes + ½ Maybe = Hit: 35/47 = .74	Automation Yes + ½ Maybe = Hit: 12/36 = .33
Experiment 2: Low (FA-Prone) Threshold				
(c) Binary				
Yes	42	16	.91	.47
No	4	18		
(d) Likelihood				
Yes	30	12	Automation Yes = Hit: 30/43 = .70	Automation Yes = Hit: 12/40 = .33
Maybe	10	10	Automation Yes + Maybe = Hit: 40/43 = .93	Automation Yes + Maybe = Hit: 22/40 = .55
No	3	18	Automation Yes + ½ Maybe = Hit: 35/43 = .81	Automation Yes + ½ Maybe = Hit: 17/40 = .43

Miss 增加
FA 减少

⇕

Miss 减少
FA 增加

Note. For the likelihood condition, we present three different means of calculating hit (H) and false alarm (FA) rate. In the first, the middle ("maybe") category is assigned to an automation "no" response; in the second, it is assigned to a "yes" response; and in the third, half of the "maybe" events are assigned to "yes" and half to "no."

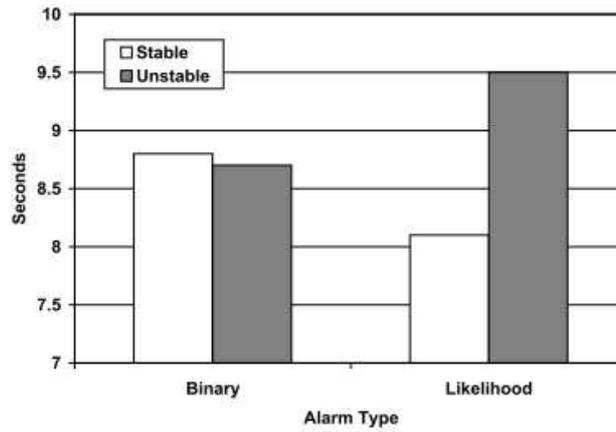


Figure 2. A significant interaction in Experiment 1 between alarm type (binary vs. likelihood) and tracking stability (stable vs. unstable) for response times (in seconds), indicating a benefit for likelihood alerting with stable tracking but a cost for likelihood alerting during unstable tracking.

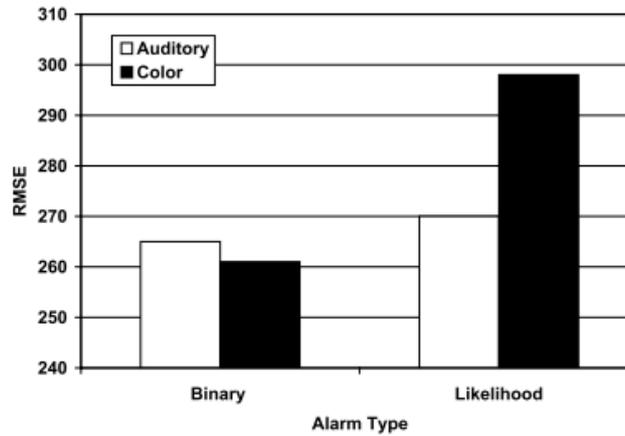


Figure 3. A significant interaction in Experiment 1 for tracking error (RMSE = root mean square error) between alert type (binary vs. likelihood) and alarm modality (auditory vs. visual) indicating a likelihood alarm cost only for visual alerts.

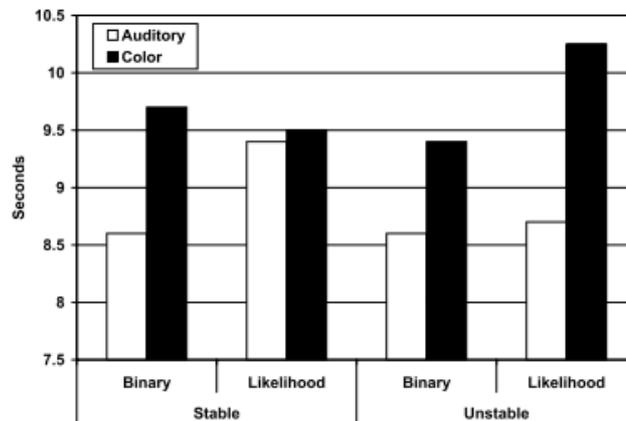


Figure 4. A significant three-way interaction for response time in Experiment 2, wherein a cost emerged for visual likelihood alerts when tracking was difficult and unstable (right) that wasn't present with the easier, stable tracking task.

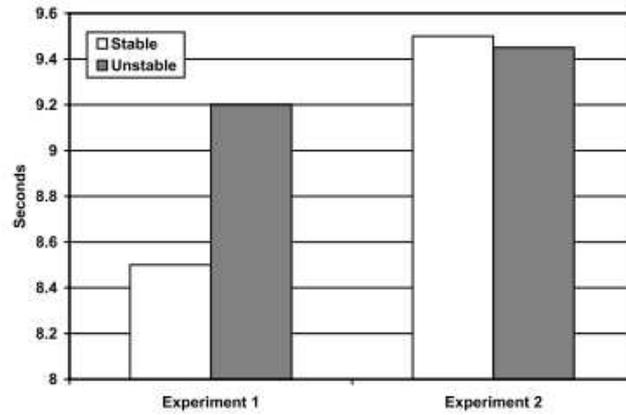


Figure 5. A significant two-way interaction for response time between alert threshold (Experiment 1, neutral, vs. Experiment 2, FA-prone) and tracking stability, such that increasing false alarm rate from Experiment 1 to Experiment 2 increased RT during stable, but not unstable, tracking.

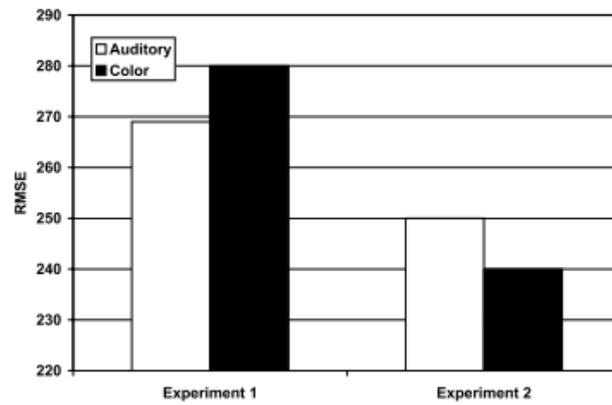


Figure 6. A significant interaction for tracking difficulty between alert threshold (Experiment 1, neutral, vs. Experiment 2, FA-prone) and alarm modality (auditory vs. visual), such that the decrease in tracking error with the decreased miss rate in Experiment 2 was greater for visual than for auditory alerts. RMSE = root mean square error.