

Examining Single- and Multiple-Process Theories of Trust in Automation

Stephen Rice

(2009). The Journal of General Psychology, 136(3), pp. 303-319.

導入

- 過去 20 年の間に自動化システムは急速に普及
 - 自動化システムのエラーが信頼や自動化システム使用に与える影響が検討されている (e.g., Bainbridge, 1982; Lee & Moray, 1994; Meyer, 2001, 2004; Parasuraman & Riley, 1997)
 - ◇ 自動化システムのエラー → 信頼低下 → 自動化システム使用の低下
 - 理論は十分ではない
 - ◇ 信頼(心理状態)と自動化システム使用(行動)
 - 完璧に一致するわけではない (Lee & See, 2004)

 - 診断システムのエラー, 信頼, 診断システムの使用との関係について検討
 - 診断システム
 - ◇ 外界の状態を推定, 診断するシステム (Alert system, Warning system)
 - ユーザの視覚的, 認知的負荷を削減
 - 診断システムを使用する上で重要な点
 - ◇ 診断システムの診断は完璧ではない
 - システムのアルゴリズムではあらゆる状況を網羅できない
 - ◇ 信号検出理論(Macmillan & Creelman, 2005)に基づくシステムエラーの解釈
 - 現実世界 : Target present, Target absent
 - 診断システムの反応 : Alert, Nonalert
- | | | |
|-----------------------|--------------|-----------------|
| | Alert | Nonalert |
| Target present | Hit | Miss |
| Target absent | FA | CR |
- ◇ 診断システムには FA と Miss の 2 種類のエラーがある
 - FA と Miss は, 診断システムへの信頼に異なる影響を与える (e.g., Dixon & Wickens, 2006; Maltz & Shinar, 2003; Meyer, 2001, 2004; Wickens & Dixon, 2007)
-
- Compliance と Reliance
 - Meyer (2001, 2004)
 - ◇ FA と Miss を区別する理論を提案

- FA はユーザの Compliance(警報に対してどう振舞うか)に影響を与える
 - Miss はユーザの Reliance(警報がない場合にどう振舞うか)に影響を与える
- FA は Compliance にのみ影響, Miss は Reliance にのみ影響 (Figure 1. B)

➤ Dixon & Wickens (2006, 2007)

◇ FA はユーザの Compliance と Reliance に影響 (Figure 1. C)

➤ Rice & McCarely (2008)

◇ FA だけでなく Miss もユーザの Compliance と Reliance に影響 (Figure 1. A, D)

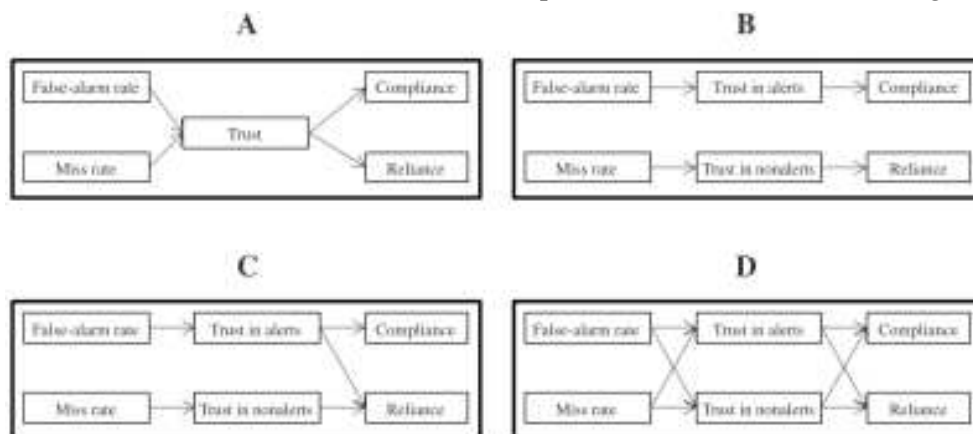


FIGURE 1. The effects of automation errors, as shown in the following theoretical models: (A) single-process model, (B) selective two-process model, (C) Mandler's two-process model, and (D) nonselective two-process model. Adapted from J. C. Dunn and K. Kirsner (1988).

- 検討を行う内容
 - ユーザは Miss と FA にどのように反応するか
 - どのモデルが最も説明力のあるモデルか

実験

- 参加者
 - 大学生 380 名
- 課題
 - ターゲット検出課題
 - ◇ 上空から敵軍のタンクを検出
 - コンピュータスクリーンに写真を 1 枚ずつ表示 (合計 100 枚)
 - ◇ 診断システムを使用しながら課題を行う
 - “The automation has detected a tank!”

- “The automation has determined that there is no tank!”

- 手順

- 診断システムの診断結果が表示される
- 5秒後に診断システムが診断を行った写真が表示される
- 写真は参加者が反応するまで表示される
 - ◇ ターゲットがあると判断 → Jキーを押す
 - ◇ ターゲットがないと判断 → Fキーを押す

- 実験計画

- エラーを起こさないシステム
- FAを起こすシステム
- Missを起こすシステム
 - ◇ 各システムの性能：55%～95%の9段階で操作
 - ◇ どの条件でも Target present の写真 50 枚, Target absent の写真 50 枚
 - FA95% : 50 hits, 5 FAs, 45 CRs
 - FA90% : 50 hits, 10 FAs, 40 CRs
 - Miss95% : 45 hits, 5 misses, 50 CRs
 - Miss90% : 40 hits, 10 misses, 50 CRs
- 全 19 条件に参加者 20 名ずつを分配して実験を実施

- 結果

- Table 1 は測定した全データ

TABLE 1. Experimental Data for Each Condition, by Reliability Level

Reliability (%)	d'		C		Compliance agreement rate		Reliance agreement rate		Compliance response time (ms)		Reliance response time (ms)		Target-present accuracy		Target-absent accuracy	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
False alarm																
95	2.53	0.80	-.23	.26	.91	.08	.90	.13	2,512.00	2,262.00	2,044.00	2,058.00	.91	.08	.83	.10
90	1.90	0.86	-.27	.41	.84	.15	.85	.18	2,079.00	1,221.00	2,099.00	1,776.00	.84	.15	.74	.14
85	1.70	0.56	-.27	.31	.85	.10	.86	.17	1,838.00	744.00	1,523.00	1,446.00	.85	.10	.71	.13
80	1.38	0.87	-.07	.35	.75	.18	.86	.21	2,280.00	1,201.00	1,540.00	1,375.00	.75	.18	.71	.15
75	1.48	0.39	-.01	.31	.77	.08	.94	.08	2,812.00	1,235.00	2,111.00	2,345.00	.77	.08	.75	.12
70	1.35	0.60	-.03	.45	.75	.10	.85	.21	2,724.00	1,388.00	1,793.00	1,855.00	.75	.10	.70	.18
65	1.33	0.76	.17	.35	.68	.10	.86	.19	2,743.00	1,453.00	2,176.00	2,242.00	.68	.10	.76	.19
60	0.95	0.59	.03	.49	.66	.15	.82	.21	3,587.00	3,140.00	2,154.00	1,893.00	.66	.15	.65	.17
55	1.21	0.80	.12	.46	.67	.14	.87	.16	2,671.00	1,153.00	1,507.00	1,584.00	.67	.14	.73	.23
Grand M	1.54	0.69	-.06	.38	.76	.12	.87	.17	2,583.00	1,532.85	1,883.00	1,841.73	.76	.12	.73	.16
Miss																
95	2.86	0.85	.30	.26	.89	.10	.83	.17	1,258.00	1,234.00	2,205.00	1,423.00	.86	.10	.93	.08
90	1.88	1.03	.25	.34	.84	.19	.85	.17	1,293.00	1,074.00	2,210.00	1,321.00	.74	.12	.83	.17
85	1.90	0.94	.22	.34	.82	.15	.76	.23	1,725.00	1,375.00	2,651.00	1,423.00	.75	.10	.83	.18
80	1.61	1.10	.19	.49	.82	.14	.80	.15	1,717.00	1,109.00	1,911.00	943.00	.72	.08	.77	.23
75	1.54	0.79	.26	.27	.82	.14	.78	.19	1,117.00	971.00	2,066.00	1,289.00	.69	.10	.81	.15
70	1.35	1.00	.24	.31	.86	.18	.78	.16	1,184.00	1,227.00	2,882.00	1,658.00	.66	.12	.77	.19
65	1.34	0.80	.28	.38	.79	.18	.78	.17	1,834.00	1,869.00	2,774.00	1,314.00	.64	.13	.78	.16
60	1.25	0.81	.27	.32	.81	.19	.68	.25	1,634.00	1,423.00	3,036.00	1,336.00	.63	.14	.78	.17
55	0.95	0.92	.08	.40	.77	.28	.66	.21	1,549.00	1,375.00	2,525.00	1,538.00	.65	.14	.68	.25
Grand M	1.63	0.92	.23	.35	.82	.17	.77	.19	1,479.00	1,295.16	2,473.00	1,360.58	.71	.12	.80	.18
100	3.80	0.22	.05	.12	.96	.08	.96	.12	356.00	138.00	401.00	124.00	.99	.04	.99	.03

➤ 敏感さ(d') (Table 2)

- ◇ d' ：検出課題における Hit と CR の正確さ
- ◇ 2(エラー：FA/Miss)×9(性能)の参加者間分散分析
 - エラー要因の主効果なし ($F(1, 342) = 1.23, n.s.$)
 - 性能要因の主効果あり ($F(8, 342) = 14.65, p < .001$)
 - システムの性能が高くなるほど検出，未検出は正確
 - 交互作用なし

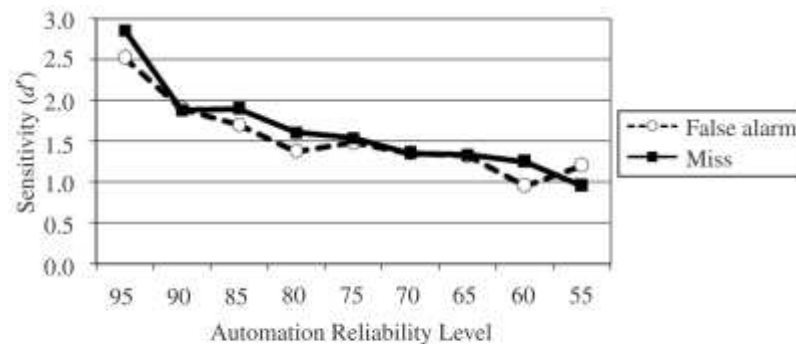
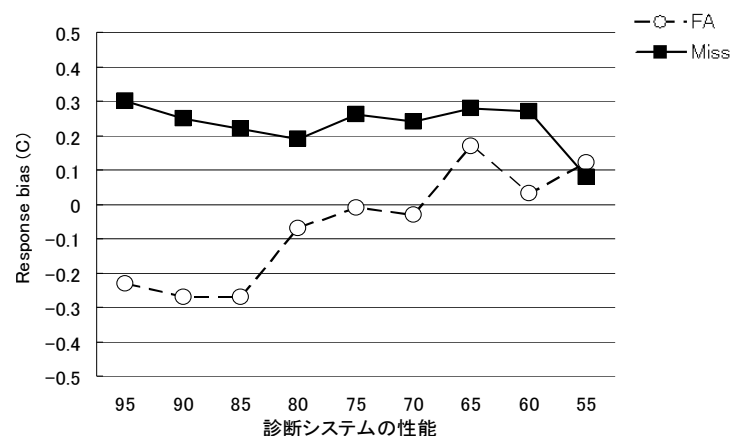


FIGURE 2. The signal-detection measure of sensitivity (d') as a function of bias and reliability.

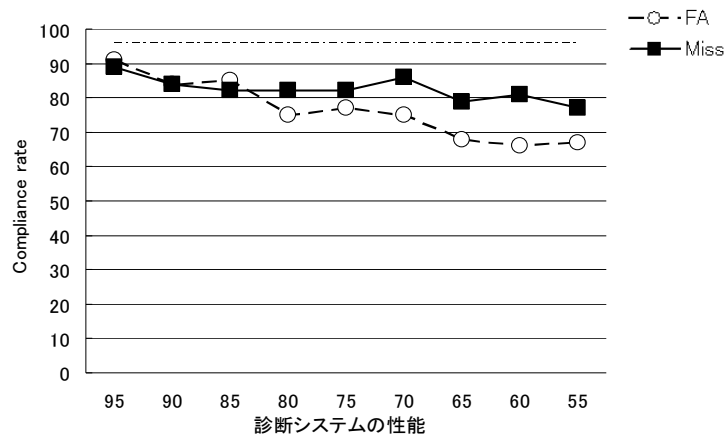
➤ 反応の偏り(C)

- ◇ C：検出課題における参加者の Target ありと判断する傾向を数値化したもの
- ◇ 2(エラー：FA/Miss)×9(性能)の参加者間分散分析
 - エラー要因の主効果あり ($F(1, 342) = 58.12, p < .001$)
 - エラーが FA の方が Target ありと判断する傾向
 - 性能要因の主効果に有意傾向($F(8, 342) = 1.87, p = .063$)
 - 交互作用あり ($F(2, 66) = 2.77, p < .01$)
 - エラーが FA の場合，反応の偏りは，性能の変化に影響を受けやすい



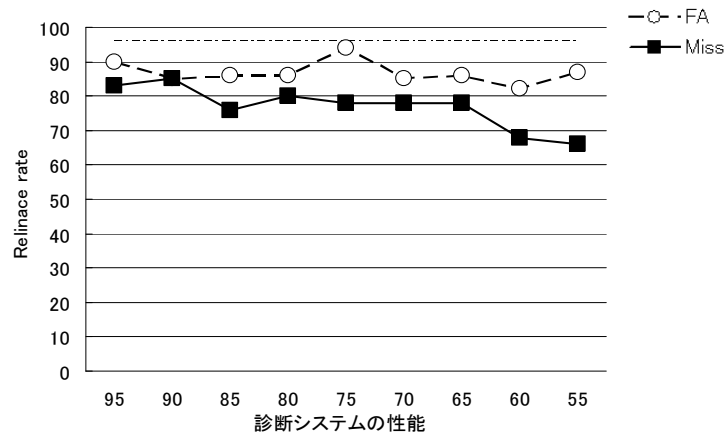
➤ Compliance

- ◇ 診断システムがターゲットありと診断し、参加者がそれに同意した割合
- ◇ 2(エラー：Miss/FA)×9(性能)の参加者間分散分析
 - エラー要因の主効果あり ($F(1, 342) = 38.81, p < .001$)
 - エラーが FA の方が同意した割合は低い
 - 性能要因の主効果あり ($F(8, 342) = 4.87, p < .001$)
 - 交互作用あり ($F(2, 66) = 2.15, p < .05$)
 - エラーが FA の場合、性能の低下に伴い、同意した割合は低下
- ◇ エラーが Miss の場合、どの条件でも、性能 100 より同意した割合は低い ($ps < .05$)



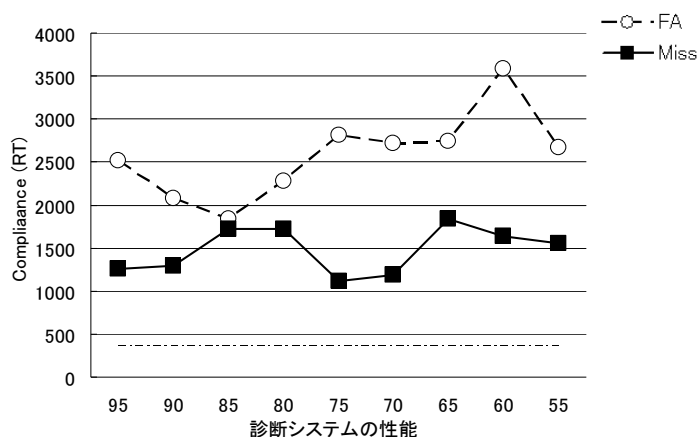
➤ Reliance

- ◇ 診断システムがターゲットなしと診断し、参加者がそれに同意した割合
- ◇ 2(エラー：Miss/FA)×9(性能)の参加者間分散分析
 - エラー要因の主効果あり ($F(1, 342) = 25.64, p < .001$)
 - エラーが Miss の方が同意した割合は低い
 - 性能要因の主効果なし ($F(8, 342) = 1.87, n.s.$)
 - 交互作用なし ($F(2, 66) = 1.14, n.s.$)
- ◇ エラーが FA の場合、どの条件でも、性能 100 より同意した割合は低い ($ps < .05$)



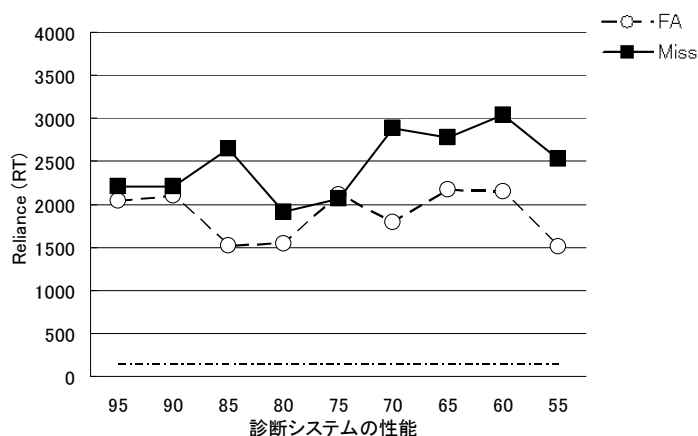
➤ Compliance の反応時間

- ◇ 診断システムがターゲットありと診断し、参加者がそれに同意したときの反応時間
- ◇ 2(エラー：Miss/FA)×9(性能)の参加者間分散分析
 - エラー要因の主効果あり ($F(1, 342) = 48.15, p < .001$)
 - エラーが FA の方が反応は遅い
 - 性能要因の主効果なし ($F(8, 342) = 1.37, n.s.$)
 - 交互作用なし ($F(2, 66) = 1.48, n.s.$)
- ◇ エラーが Miss の場合、どの条件でも、性能 100 より反応は遅い ($ps < .05$)



➤ Reliance の反応時間

- ◇ 診断システムがターゲットなしと診断し、参加者がそれに同意したときの反応時間
- ◇ 2(エラー：Miss/FA)×9(性能)の参加者間分散分析
 - エラー要因の主効果あり ($F(1, 342) = 33.70, p < .001$)
 - エラーが Miss の方が反応は遅い
 - 性能要因の主効果なし ($F(8, 342) = 1.57, n.s.$)
 - 交互作用なし ($F(2, 66) < 1.0, n.s.$)
- ◇ エラーが FA の場合、どの条件でも、性能 100 より反応は遅い ($ps < .05$)



➤ State trace analyses

- ◇ FA と Miss で各性能が Compliance と Reliance に与える影響について検討
- ◇ FA と Miss の各性能における Compliance と Reliance の割合, 反応時間の散布図 (Figure 3)
 - 相関なし : 割合 ($r = -.28, n.s.$), 反応時間 ($r = -.35, n.s.$)
 - FA の場合, 性能の違いによって Compliance への影響が異なる
 - FA は Reliance よりも Compliance に影響
 - Miss の場合, 性能の違いによって Reliance への影響が異なる
 - Miss は Compliance よりも Reliance に影響
- ◇ このようなパターンは, Figure 1A のモデルでは説明できない
 - 単一の信頼機能に基づいて Miss, FA を評価
 - Compliance と Reliance に相関関係がみられるはず

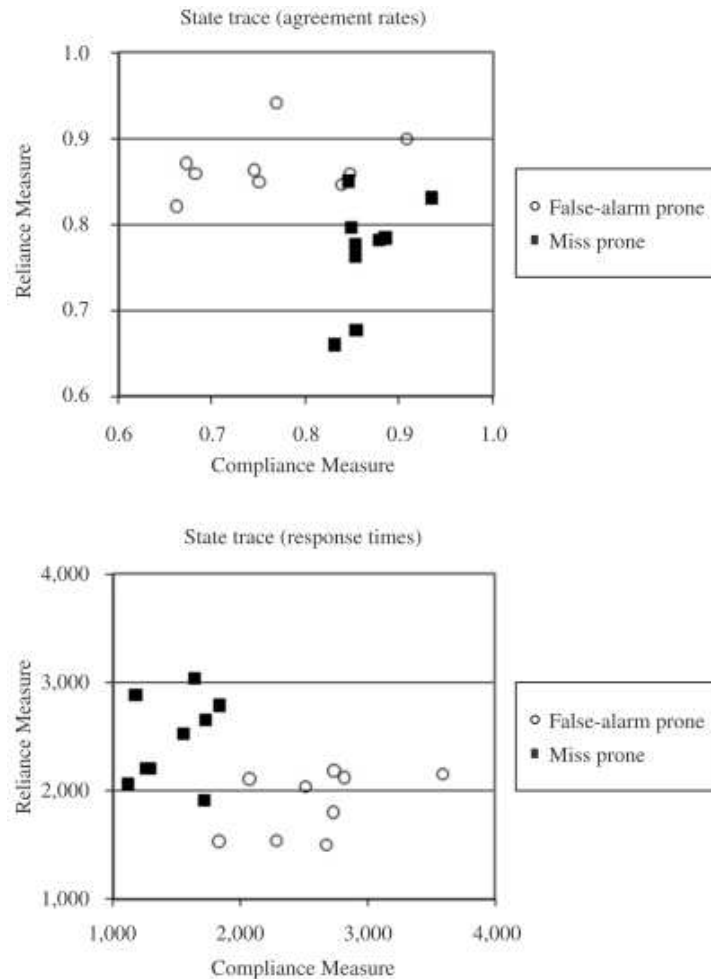


FIGURE 3. State-trace analyses on agreement rates (%) and response times (ms). Solid squares indicate miss-prone conditions, whereas clear circles indicate false-alarm-prone conditions.

考察

● Figure 1B のモデル

- 今回の実験結果からこのモデルは支持されない
- ◇ FA は Compliance と Reliance に影響
 - Miss よりも Compliance の割合は低く，反応時間は遅い
 - 性能 100% よりも Reliance の割合は低く，反応時間は遅い

● Figure 1C のモデル

- 今回の実験結果からこのモデルは支持されない
- ◇ Miss も Compliance と Reliance に影響
 - FA よりも Reliance の割合は低く，反応時間は遅い
 - 性能 100% よりも Compliance の割合は低く，反応時間は遅い

● Figure 1D のモデル

- 今回の実験結果から支持されるモデル
- ◇ FA は Reliance よりも Compliance に影響を与える
- ◇ Miss は Compliance よりも Reliance に影響を与える
 - Figure 1A のモデルでは，同程度に影響するはず