False Alerts in Air Traffic Control Conflict Alerting System: Is There a "Cry Wolf" Effect?

Christopher D. Wickens, Alion Science Corporation, Boulder, Colorado, **Stephen Rice** and **David Keller,** New Mexico State University, Las Cruces, New Mexico, **Shaun Hutchins,** Alion Science Corporation, and **Jamie Hughes** and **Krisstal Clayton,** New Mexico State University, Las Cruces, New Mexico

Objective: The aim is to establish the extent to which the high false-alarm rate of air traffic control midair conflict alerts is responsible for a "cry wolf" effect—where true alerts are not responded to and all alerts are delayed in their response. **Background:** Some aircraft collisions have been partly attributed to the cry wolf effect, and in other domains (health care and systems monitoring), there is a causal connection between false-alarm rate and cry wolf behavior. We hypothesized that a corresponding relationship exists in air traffic control (ATC). **Method:** Aircraft track and alert system behavior data surrounding 495 conflict alerts were analyzed to identify true and false alerts, trajectory type, and controller behavior. Forty-five percent of the alerts were false, ranging from 0.28 to 0.58. **Results:** Although centers with more false alerts contributed to more nonresponses, there was no evidence that these were nonresponses to true alerts or that response times were delayed in those centers. Instead, controllers showed desirable anticipatory behavior by issuing trajectory changes prior to the alert. Those trajectory pairs whose conflicts were more difficult to visualize induced more reliance on, and less compliance with, the alerting system. **Conclusion:** The high false-alarm rate does not appear to induce cry wolf behavior in the context of en route ATC conflict alerts. **Application:** There is no need to substantially modify conflict alert algorithms, but the conflict alert system may be modified to address difficult-to-visualize conflicts.

INTRODUCTION

In June 2006, it was recommended by the Federal Aviation Administration (FAA) that the FAA research community "consider how and when an alert is present, and offer solutions to improving this process" (FAA, 2006). Also in 2006, the National Transportation Safety Board (NTSB) advised the FAA to "redesign the Conflict Alerting system . . . to reliably direct controller attention to potentially hazardous situations detected by the system" (NTSB, 2006). Both documents made reference to a series of accidents in which the minimum safe altitude warning (MSAW) and conflict alerts (CAs) announced pending controlled (flight-into-terrain [collisions]) and midair collisions, respectively. However, controllers failed to respond or intervene to prevent the accidents.

Furthermore, anecdotal evidence from a specific accident (midair collision of two aircraft near San Diego), and from other interviews with controllers (Ahlstrom & Panjwani, 2003), suggested the prevalence of controller experience of the "cry wolf" effect (Breznitz, 1983; Sorkin, 1989). The NTSB (2006) report stated that "controllers repeatedly cited the number of unwarranted 'nuisance alarms' that they are exposed to on a routine basis" (p. 7). In addition, the report also stated that "alarms that go off too frequently, especially false alarms" (FAs; NTSB, 2006, p. 7) is among the five most commonly expressed issues with alarms. The "cry wolf" effect is a general syndrome whereby

Address correspondence to Christopher D. Wickens, Alion Science Corporation, Micro Analysis and Design, 4949 Pearl East Circle, Suite 300, Boulder, CO 80301; cwickens@alionscience.com. *HUMAN FACTORS,* Vol. 51, No. 4, August 2009, pp. 446-462. DOI: 10.1177/0018720809344720.

excessive alarms, many of them seemingly unnecessary to the operator (e.g., FAs or false alerts), lead to a distrust, or disuse, in the alarm system. In turn, this operator distrust, or disuse, leads to a disregard of (or late response to) some true alarms (Lee & See, 2004; Parasuraman & Riley, 1997). Linking this well-observed phenomenon to the findings of missed alerts in the NTSB (2006) study suggests that there may be a causal connection between the cry wolf behavior and FAs.

The supporting research on imperfect (e.g., FA-prone) conflict alerting can be approached from two different perspectives in the context of signal detection theory (SDT), corresponding to the two SDT parameters of sensitivity (d') and response bias (beta) as applied to alerting system performance (Getty, Swets, Pickett, & Gonthier, 1995; Swets & Pickett, 1982). The approach based on d' asks how low the reliability of an alerting or diagnostic system can get while preserving performance that is superior to that of the unaided human (Wickens & Dixon, 2007; Xu Rantanen & Wickens, 2007). The approach based on beta, or the "threshold" of the alerting system (the direct intent of the current research), addresses the trade-offs between a high threshold of the alerting system, which creates system misses, and a low threshold, which creates a higher rate of cry wolf–inducing FAs.

A general conclusion from research in this area appears to be that although misses may be catastrophic in a multitask system in which there is no human backup to monitor the raw data in parallel, in systems that allow such parallel human–machine monitoring (Getty et al., 1995; Parasuraman, 1987), FA-prone systems may often be worse (Dixon & Wickens, 2006; Dixon, Wickens, & McCarley, 2007; Maltz & Shinar, 2003; see Wickens, Levinthal, & Rice, in press, for a summary). This conclusion may be particularly true in high-workload multitask circumstances, given that this response can be quite disruptive to concurrent tasks either as a result of carrying out the unnecessary alert-triggered action (Stanton & Babar, 1995) or as a result of the need to cross-check the raw data to establish that the alert was indeed false. In a further argument for a higher threshold, in many *predictive* alerting systems, such as the CAs studied here, a higher threshold translates not necessarily to more missed events but only to a later alerting of true events (a much less catastrophic outcome than a miss) (Kuchar & Young, 2000). Indeed, if this alerting look-ahead time still provides adequate time for humans to respond, then the benefit of reducing FAs would more than offset the cost of the shorter period between the alert and the occurrence of the forecast event (e.g., the pending collision).

To further complicate the picture, Lees and Lee (2007) have introduced the distinction between truly "bad" false alerts and "acceptable" false alerts. In the context of their car-driving headway monitoring study, bad false alerts occur at random times, unrelated to the state of the raw data (in their study, visual view of headway). In contrast, acceptable false alerts occur when the headway trend is shortening but appear to be premature, as if the alerting system has merely set too low a threshold (e.g., the designer's choice to err on the side of decreasing late alerts). Thus, to the extent that false alerts are "acceptable," they can be viewed as actually helpful in confirming to humans that the alert system is generally working well. Lees and Lee confirmed the benefit of these acceptable false alerts, just as they confirmed the harm (mediated by cry wolf) of the random-appearing FAs.

A related concept that we invoke in the current study is the nature of *anticipatory behavior*, whereby a human acts in response to a dangerous event prior to an alert event (Levinthal & Wickens, 2005; Meyer & Bitan, 2002; Wickens et al., in press; Woods, 1995). In this case, it is possible to think of either the alert as delayed (the result of a high threshold) or the human as particularly vigilant, or proactive, in monitoring the raw data, demonstrating desirable anticipatory behavior (Burns, 2006).

We can consider this anticipatory behavior, observed by Levinthal and Wickens (2006), and the nonresponsive behavior, evident in the cry wolf phenomenon, as different manifestations of the two aspects of *automation dependence*—reliance and compliance—introduced by Meyer (2001, 2004; also see Dixon et al., 2007; Dixon & Wickens, 2006; Maltz & Shinar, 2003; Rice, in press). *Reliance* describes circumstances in which the human does not respond when the alert system is "silent." Hence, anticipatory behavior indicates a lack of reliance. *Compliance* describes circumstances in which the human does respond when the alert occurs. Hence, the nonresponse or delayed response to the alert that characterizes the cry wolf effect describes a lack of compliance.

The purpose of the current study was to seek evidence for the FA-caused cry wolf phenomenon from live, or "naturalistic," data across five air traffic control (ATC) facilities. Within these five facilities, controllers responded to varied midair CAs. Such live ATC data have never before been examined in this fashion, although it parallels the analysis of professional weather forecasters (Barnes et al., 2006), pilots (Bliss, 2003), and health care practitioners (Lawless, 1994; Vashitz et al., 2008; Xiao, Seagull, Nieves-Khouw, Barczak, & Perkins, 2004) responding to imperfect alerting systems. In this process, we must first examine performance of the CA system itself, assess an FA rate, and then examine the influence of differences in this rate on behavior of the controller and on performance of the controller-CA (human-automation) system as a whole.

In this study, we also distinguish between two aspects of FA-induced behavior. With regard to the first aspect, there is behavior triggered by the event that produced a false alert itself—behavior that may be related to the nature of the event that made the alert false. For example, it may be a particularly challenging event that caused the alerting system to make an error (Madhavan, Wiegmann, & Lacson, 2006) or the controller to respond inappropriately. With regard to the second aspect, there is the "proneness" of the system to make false alerts or missed alerts across a series of conflict events (lowering its reliability). This can induce a cognitive set of distrust or "low compliance" (Meyer, 2001, 2004) that can express itself as a cry wolf syndrome on alerts that are either true or false. Naturally, of greatest concern is the cry wolf delayed response or nonresponse for events when the alert is true.

In the current research, we address three primary hypotheses. First, we predict there will be a number of false alerts in the data and some variance across this FA rate. In essence, because this is not a controlled experiment, this hypothesis is that we will find an "active" independent variable—FA rate—that can be used to test our second hypothesis. Our second hypothesis can be assessed in two forms: (a) The existence of a substantial number of false alerts will produce nonresponses to true alerts. (b) ATC centers that have a higher false-alert rate will show greater evidence for cry wolf behavior (i.e., later responses and/or more nonresponses). Third, we predict that (a) reliance on automation may be reduced when conflicts are easier to visualize in the raw data, because they may be anticipated prior to the alert, and (b) compliance with automation may also be increased by ease of visualization, because it is easy to see that the alert will be true.

Because the current study was based on data provided to us by the FAA, and we were requested to perform the analysis after the data had been collected, we had neither the opportunity to collect specific data (e.g., trust ratings) or access the demographics of controllers involved (e.g., level of experience); controller information was kept confidential. Therefore, we were unable to infer causation with high confidence.

METHOD

The CA System

The CA system (FAA, 2003) (Paglione, Ryan, & Liu, 2007) is designed to predict when two aircraft are within 5 miles laterally and 1,000 ft vertically of one another (see Figure 1). Such closure is known as a *loss of separation* (LOS). Hence, the CA predicts any LOS that is forecast to occur within 75 to 135 s. The system samples enough radar data to make a stable extrapolation of the trajectories—an amount that will show variation from pair to pair. Because of the range of radar data quality, there will be a range of times required to obtain a stable estimate, the source of the 60-s range of look-ahead values. When the CA system predicts such an LOS, the data tags flash on the controller's display. In the en route centers where performance was evaluated, there is no auditory sound associated with the CA.

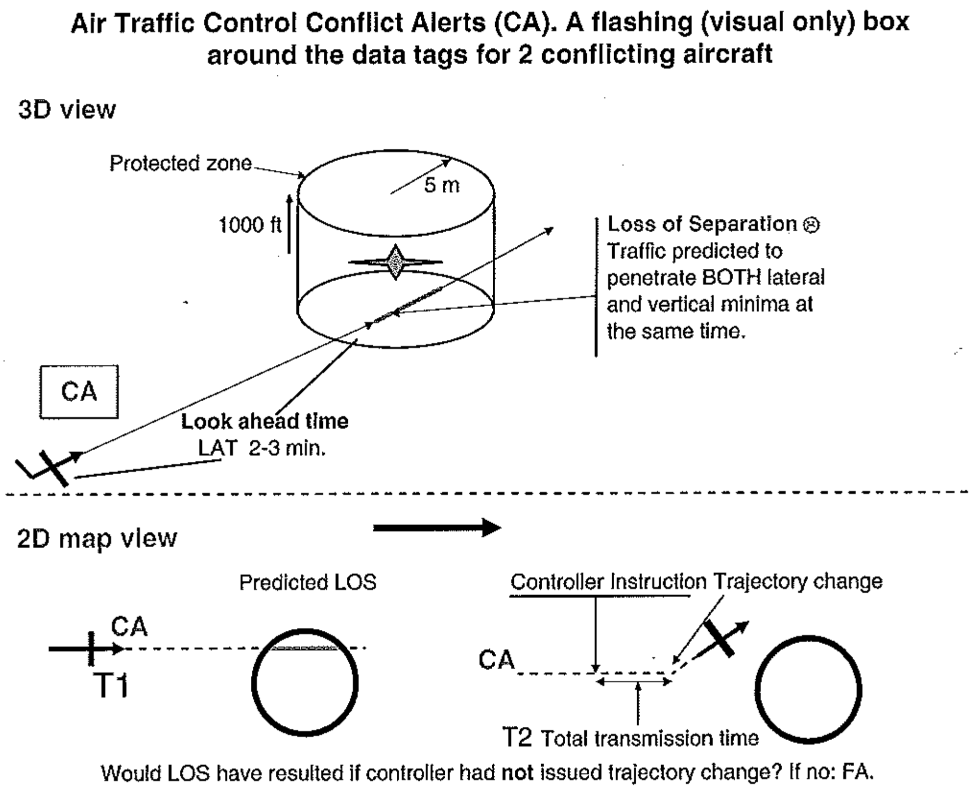

*Figure 1.*Schematic representation of the conflict alert system. A 3-D version at the top, two examples of a 2-D top-down version at the bottom: On the left, there is a predicted loss of separation (LOS; conflict). On the right, the controller has initiated a trajectory change before the loss of separation occurs. FA = false alarm.

The algorithm underlying the CA generates a linear extrapolation on both the horizontal (map) plane and the vertical plane of the current heading and vertical speed of both aircraft, respectively (FAA, 2003). The algorithm contains no knowledge of intent (e.g., no representation of a flight plan that may later trigger a leveling off of one aircraft or the other). Because the air traffic controller has unique knowledge of traffic, the pilot must comply with the air traffic controller's instructions following a CA. The only instance in which the pilot should not comply is when the pilot senses a direct danger to the aircraft as the result of such compliance.

Data

We were provided data from the FAA for 495 conflict alerts, extracted from the busiest 2-hr periods from a sample of 2 or 3 days in each of five en route ATC centers. Such data included, for each CA, properties of the pair of trajectories predicted by the CA (e.g., predicted point of closest passage, time of alert), the actual radar tracks and altitude of the aircraft (sampled every 10 s), and a short analysis of the actual conflict as it was played out (see Wickens et al., 2008, for details). The most important element of this third set was a metric (min-max ratio, or MMR) describing the severity of the conflict. Because an LOS is a simultaneous reduction of separation below the criteria of 1,000 vertical feet and 5 miles, the MMR scales each of these two values relative to 1 (so 1.0 = the criterion value, and 0 means no separation on the axis in question). For each CA, the MMR algorithm reports the maximum of these two minimum ratios. A value of 0 corresponded to an actual collision. A value of 1 was the threshold for an LOS. Progressively higher values above 1.0 indicated passage with greater lateral and vertical separation than the minima.

TABLE 1: Basic Data From Conflict Alert (CA) Systems

Variable	ZLC	ZHU	ZLA	ZTL	ZID
Encounters per hour[a]	1,126	1,589	5,529	5,679	8,813
No. of CAs[b]	22	36	72	235	124
CA rate[c]	22/4,525 = 0.005	36/4,767 = 0.007	72/16,589 = 0.004	235/38,815 = 0.006	124/26,440 = 0.004

a. Number of encounters per hour as an estimate of busyness of the center.
b. Total number of conflicts from the center that were provided to the authors (for a differing number of days, depending on the center).
c. An estimate of CA rate, reached by dividing the number of conflicts by the number of encounters, measured during the hours for which data were provided.

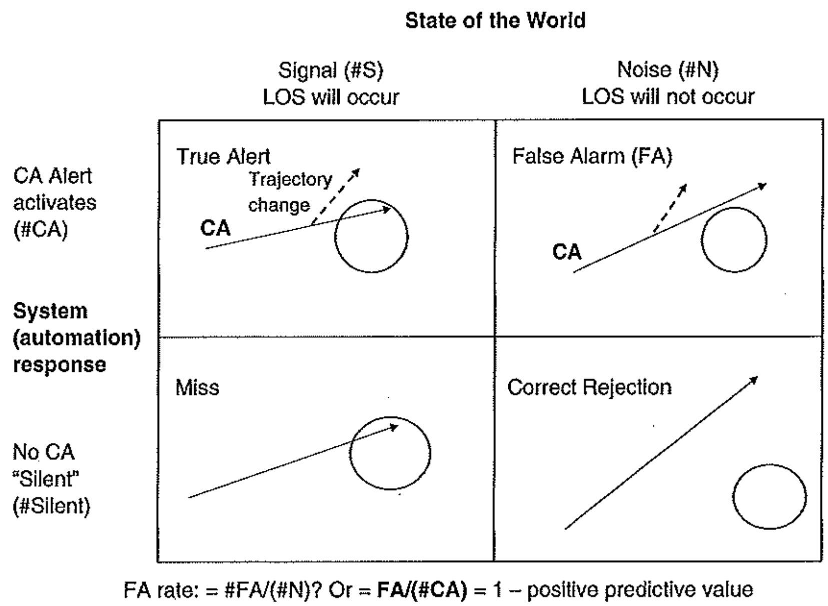

Figure 2. Signal detection matrix for conflict alerting. Within each cell is depicted the predicted trajectory (in the horizontal plane only) relative to the protected zone of another aircraft. In the top two cells are represented the two cases in which the trajectory either does (dashed line) or does not (solid line) change prior to the point of closest approach.

Two key variables we derived from the data for each center were the "busyness" of the center (the number of encounters per hour, in which "encounter" is an instance in at which the CA algorithm begins to examine track pairs; 40 miles, 5,000 vertical feet) and the total number of CAs provided to us for analysis. Table 1 shows these two parameters across the five centers (rows 1 and 2) along with, in row 3, the ratio of the total CAs to the total encounters within the center during the equivalent period—an estimate of the CA rate. Importantly, Table 1 reveals that what can be defined as the "CA rate" in the bottom row did not vary substantially across centers, a mean value of 0.0057.

CA System Analysis

The typical signal detection matrix for a diagnostic system is represented as shown in

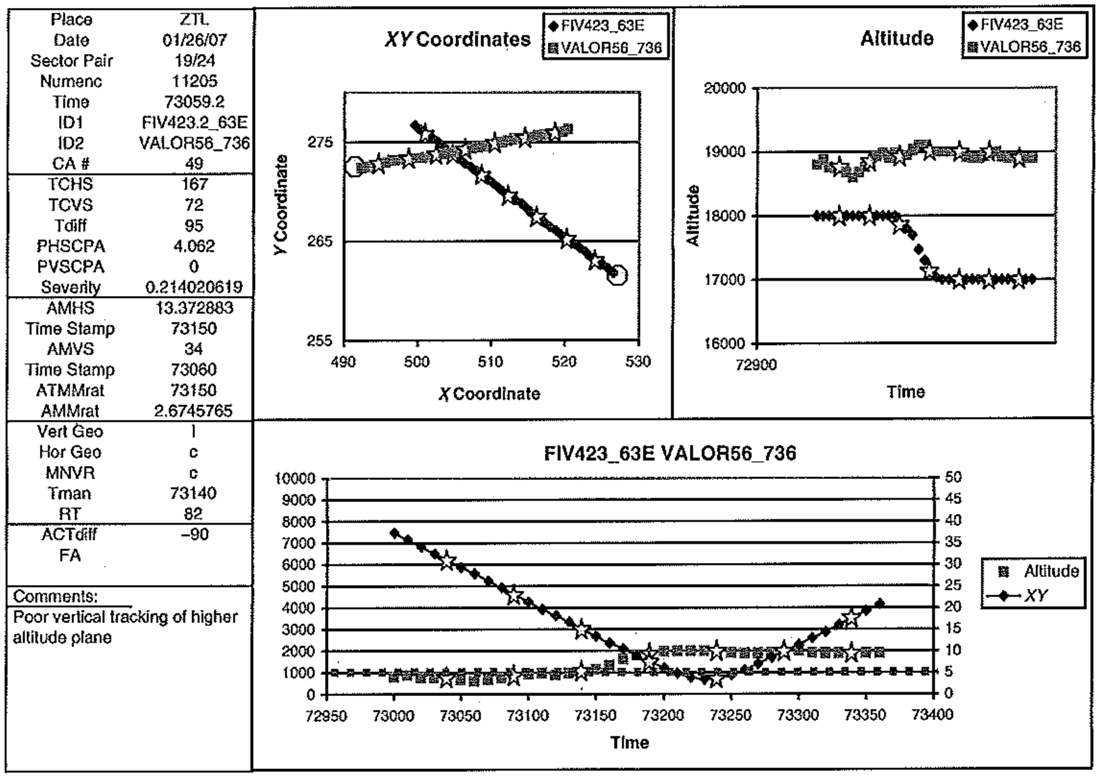

Figure 3. Example of the raw trajectory data, which were rendered visually. In the bottom panel, the altitude separation (gray squares) is flat, remaining below 1,000 ft until a correction is implemented at time 73120, whereas the XY separation (black diamonds) converges gradually and then diverges. Further explanation of the elements is contained within the text.

Figure 2, with a false-alert rate defined as the number of false alerts divided by the number of times that an LOS was not predicted to occur between aircraft pairs. However, as shown in the figure, this computation was challenged in two respects. First, as represented in the top two cells of the matrix, on some occasions, the controllers responded to the CA by altering trajectory (heading and/or vertical speed) of one of the aircraft in a pair (the dashed lines), and on other occasions, no trajectory change was initiated. In the latter case (the solid line), it was easy to determine whether a CA was true or false. A true alert would be one that produced an LOS. In the former case, when a trajectory was altered by an instruction from the controller (the dashed line), it was necessary to create a graph of the track data prior to the alteration. Such a graph is shown in Figure 3.

Figure 3 provides an example of the raw trajectory data, which we rendered visually. On the top left we show the map (*XY*) track of the two aircraft that triggered the CA. On the top right we have graphed the vertical profile. Importantly, from the map plot, there was no turn evident, and in the vertical profile, there was a clearly defined descend maneuver for the lower aircraft, a maneuver whose time following the CA could be specified. At the bottom, we present the separation graph along the two axes. The increase in vertical separation (altitude difference), following the descend maneuver of the lower aircraft, is clearly evident. To assess that an LOS would have occurred had the aircraft not been maneuvered, it is necessary to extrapolate this vertical separation curve forward to establish that vertical separation would have remained below minimum at the time that lateral separation was also lost, as shown by the negative parabola curve of *XY*. Because, however, the corrective trajectory change was made, there was not an LOS on both axes at the same time. Hence, an LOS was not observed, but the trial was categorized as a true alert.

The extrapolation of the separation functions in the figure to determine whether the trajectory pair would have simultaneously compromised lateral and vertical minima, had the alteration not taken place, was carried out by two independent observers for two of the centers and by one of these observers for the remaining three. We analyzed the data in two batches, 1 year apart. The first batch (two centers) included both observers. The second observer was not available for the second batch. Because there was high agreement between observers for the first batch, and the skilled observer was available for both, we chose not to train another observer.

It also became meaningful to examine separately properties of CAs when controllers did and did not respond (as signaled from the presence or absence of a trajectory change). This is discussed later.

The second challenge resulted because we did not have track data for any of the pairs that did not trigger a CA (e.g., the data in the bottom two cells of Figure 2). Hence, it was impossible to assess the number of noise trials or the other properties of these "encounter but no CA" trials. As a consequence of the second challenge, we redefined FA rate to be that proportion of CAs that were false. This ratio is equivalent to the inverse of what is called the predicted positive response (1 – proportion of CAs that are true). In the following, we refer to this measure as the FA rate.

RESULTS

CA System Analysis

There was no relationship between CA rate and FA rate. As discussed in more detail below, FA rate did vary considerably between 0.28 and 0.58. On the average approximately half of the CA's are false, a value that is substantial, although a good deal lower than in some systems, where the base rate is very low (Getty et al, 1995; Krois, 1999). There was no relation between CA rate and FA rate, and only a weak correlation was created almost entirely by a single outlier point from the lowest density center. These results satisfied our first hypothesis—there was substantial variance in FA rates between centers.

We also analyzed and categorized the geometry of the trajectories of the pairs of aircraft entering into each CA. Horizontal geometry was classified in terms of whether the pair was converging (C), diverging (D), or parallel (and heading in the same direction [P]; parallel tracks in opposite directions were categorized as converging). Vertical geometry was broken into five major categories: both level (L), one level and one nonlevel (N), both nonlevel and converging (C, e.g., climbing and descending toward each other), parallel climb or parallel descent (P), and both nonlevel and diverging (D). Note that a CA can be triggered even as the pair is diverging on one axis, as long is there is more rapid convergence on the other axis so that the planes are predicted to go below minimum separation on one axis before they go above minimum on the other.

Controller Performance Analysis

Categorical analyses. Before examining the relation between the change in FA rate and manifestations of the cry wolf phenomenon (Hypothesis 2b), it was necessary to identify the overall prevalence of those manifestations in our sample of data (Hypothesis 2a) Thus, in addition to the dichotomization of true versus false alerts discussed earlier, we examined two other important dichotomies that are characteristic of the human (controller): (a) the presence or absence of a response (as inferred from visual analysis of the track data discussed earlier) and (b) the presence or absence of an LOS, as revealed by MMR < 1.0.

Overall, it was usually relatively easy to identify whether a distinct change in trajectory was implemented in the time period following a CA, allowing inference of a controller response (see Figure 3). However, for a small sample, this classification became quite difficult because of jitter in the radar track data (particularly, vertical). Therefore, those trials were not included in the data base for analysis. We note also in Tables 2 through 5 that the total N for different classifications was not always equivalent. This is because a classification for CA for one dichotomous variable may be uncertain (and hence discarded) but retained for another.

TABLE 2: Frequency of True (TA) and False Alerts (FA) for Responses and Nonresponses

Alert Type	No Response	Response	Total
TA	3	231	234
FA	37	166	203
Total	40	397	437

TABLE 3: Frequency of True (TA) and False Alerts (FA) for Loss-of-Separation (LOS) and Non-LOS Trials

Alert Type	No LOS	LOS	Total
TA	164	34	198
FA	170	5	175
Total	334	39	373

Table 2 shows the frequency of true versus false alerts that produced either a controller response or no controller response.

Table 2 reveals that for roughly 10% of the CAs, there was no evidence for a controller response, at least as indicated by a trajectory change by either of the two aircraft involved in the CA. These nonresponses were statistically more prevalent when the CA was false (18%) than when it was true, 1.5%, $\chi^2(1, N = 437) = 37.5$, $p < .01$. Such a result is quite plausible, given that the trajectories triggering a false alert are, by definition, more likely to yield a more distant "closest passage" or miss distance and, in turn, more likely to be considered by the controllers not to require a trajectory change.

Table 3 shows the frequency of an LOS as a function of whether the alert was true or false. Note that the total LOS is much smaller than in Table 2. This is because, for several conflicts, the measure MMR, which allowed classification of LOS (<1.0), was not provided in the data base given to us. Table 3 reveals that the LOS rate is, like the nonresponse rate, approximately 10% of the data base. Also, it appears that the two types of outcomes are unevenly distributed across the two types of alerts. Specifically, Table 3 demonstrates that true alerts are more likely to precede an LOS (21%) than are false alerts, 3%; $\chi^2(1, N = 373) = 20.3$, $p < .0001$. This is a plausible outcome, given that the true alert will occur on a trajectory pair that is more dangerous and, therefore, slightly more likely to yield the ultimate LOS even following a controller intervention.

In Table 4, we cross controller response and LOS to establish the extent to which controller nonresponses might be associated with an LOS. These observations are collapsed across true versus false alerts. The data in Table 4 indicate that when the controller did not respond, this

TABLE 4: Frequency of Controller Response for Loss-of-Separation (LOS) and Non-LOS Trials

Trial Type	No Response	Response	Total
LOS	2	37	39
No LOS	32	309	341
Total	34	346	380

was very unlikely to produce an LOS (5%; and those two events were restricted to a single center), whereas when the controller did respond, such LOS events were somewhat more prevalent (9%) although the difference in proportion was not significant, $\chi^2(1, N = 380) = .778$, $p > .10$. We note here that this finding does not necessarily imply that controller responses were counterproductive, because as Table 4 suggests, the vast majority of LOS cases occur on true alerts. In these, had the controller not intervened with a trajectory change, there definitely would have been an LOS.

Collectively, the previous three analyses provide no evidence for the most dangerous manifestation of the cry wolf effect (nonresponse leading to a LOS as a opposed to a nonresponse leading to no LOS). The number (2) of such joint events is even fewer than what the independent product of the two classes of events might predict (10% nonresponse rate × 10% LOS rate = 1% of the CA events = 4). Therefore, Hypothesis 2a was not confirmed.

To explicitly test Hypothesis 2b, we next sought to determine whether there was any relationship between FA rate (as it varied across centers) and either nonresponses or LOS events (although we note that in the absence of controlled manipulation, true causality is hard to establish with certainty). Figure 4 shows the

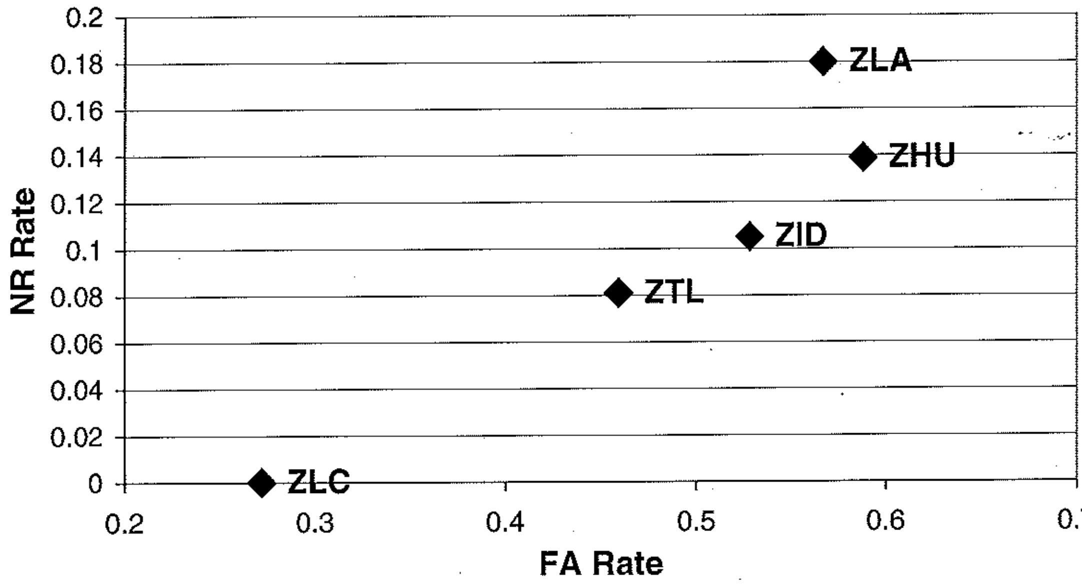

Figure 4. Nonresponse rate (NR) as a function of false-alarm (FA) rate across centers.

scatter plot of FAs versus nonresponses and reveals a striking and pronounced trend: The greater the false alert rate in the center, the less controllers tended to respond, $r = .944$, $p < .05$. However, when the LOS rate was examined as a function of FA rate across center, there was no trend. This null effect suggests that the increase in nonresponses in the more FA-prone center, shown in Figure 4, were not associated with less safe separation.

Response time (RT). We then sought evidence for the second manifestation of the cry wolf phenomenon—the possible delay in RT associated with more FAs. Interpreting the delay between the CA and the trajectory change response required consideration of the total transmission lag (TTL), which is the time for the following processes to occur: First, controller notices a dangerous convergence; second, controller chooses a trajectory change and communicates this to the pilot; third, pilot confirms and implements the change with the flight controls; and fourth, the aircraft initiates a sufficient trajectory change to be evident in the radar track. This TTL is estimated to be approximately 20 s to 25 s on the basis of voice transcript analysis (Allendoerfer & Friedman-Berg, 2007; Friedman-Berg, Allendoerfer, & Pai, 2008), analyses that were not available for the CA trials examined in the current study.

Our analysis revealed that for about 45% of the CAs, controllers must have initiated the perceptual and cognitive trajectory processing (noticing convergence and choosing a maneuver) before the alert occurred, because in these trials, the measured RT was less than 25 s. Indeed, when we examined the distribution of RTs, relative to the CA, we observed a bimodality, shown in Figure 5, with a local minimum at approximately 25 s. This bimodality, coupled with the estimate of a 25-s TTL, supported the notion that there were two categorically different types of responses: anticipatory responses and reactive responses.

The RT data were positively skewed, so a log transformation was carried out to reduce the skew. An ANOVA carried out on the log-transformed RT data indicated that for anticipatory responses, there was no difference in RT between true and false alerts ($p > .10$); however, for reactive responses, true alerts were responded to approximately 14 s ($73 - 59 = 14$) more rapidly than were false alerts, $t(193) = 2.4$, $p < .02$, reflecting the increased urgency of the former. There was no significant difference in RT between LOS and non-LOS encounters, so

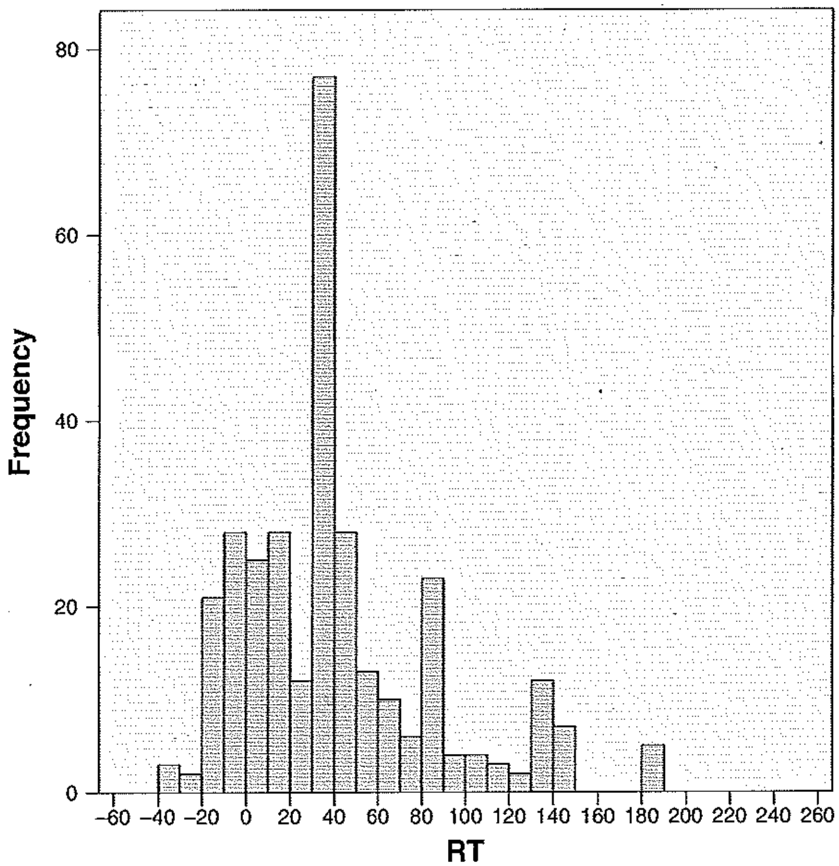

Figure 5. Distribution of response times (RT; in seconds) relative to the time of appearance of the conflict alert.

an inference that the LOS resulted because of a delay in responding was rejected.

An analysis of three centers' data did reveal a main effect of center, $F(2, 154) = 6.78, p < .01$, with the highest-density center (ZID) showing faster responses (30 s) than either the low-density (ZHU, 33 s) or mid-density (ZLA, 36 s) centers. This effect was observed for both anticipatory and reactive responses. This effect is noteworthy because whereas increasing density might have been anticipated to slow RT because of greater workload, the faster RT for ZID was observed despite its greater traffic density (see Table 1).

Finally, within the non-LOS encounters (MMR > 1.0), we correlated RT with the value of MMR to test if later-responses trials produced closer (but still above minima) passages. This correlation, examined with the pooled data for the three midlevel centers, was not significant ($p > .10$). This finding suggests that controllers did not compromise safety when their responses were delayed.

TABLE 5: Anticipatory Versus Reactive Responses for True (TA) and False Alerts (FA)

Alert Type	Anticipatory	Reactive	Total
TA	109	114	223
FA	56	95	151
Total	165	209	374

Categorical analysis of anticipatory response frequency. Table 5 depicts the frequency of anticipatory versus reactive responses for true versus false alerts. Analyses of these data reveals that controllers were significantly more likely to anticipate on a true (0.58) than on a false (0.37) alert, $\chi^2(1, N = 374) = 5.08, p < .05$.

Table 6: Categorical events as a function of lateral geometry. Measures in proportion (P) of events

Variable	Parallel	Converging	Diverging
Lateral geometry, false alert (FA) rate			
P(FA)	0.45	0.47	0.36
n	80	321	25
Lateral geometry, anticipation			
P(anticipation)	0.30	0.51	0.11
n	71	299	18
Lateral geometry, response			
P(response)	0.88	0.93	0.60
n	82	330	30

This is a plausible finding because the true-alert trajectories should signal the impending conflict with greater salience in the raw data of the radar displays.

Trajectory analyses. Final analyses were undertaken to assess whether particular trajectory combinations of the two planes were associated with more FAs, reactive responses, or nonresponses—all characteristics that would suggest that certain trajectories were harder to process than others. The frequency of events as a joint function of the three horizontal geometry categories and alert type (true vs. false) is shown in Table 6 for each of the three categorical variables expressed as the following proportions: P(FA), P(anticipation) and P(controller response). The number of observations (denominator of the proportion) is shown underneath each proportion. We note here, as in Tables 2 through 5, that the total number of observations is not always the same across the analyses. A chi-square analysis on the FA rate in Table 6 revealed that the false-alert rate was independent of lateral geometry, $\chi^2(2, N = 426) = 1.26, p > .10$.

Table 6 also presents the frequency of anticipatory responses as a function of lateral geometry. The significant chi-squared test, $\chi^2(2, N = 388) = 19.3, p < .01$, revealed that anticipatory responses were much more likely when tracks were converging on the radar display (51%) than when they were either parallel (30%) or diverging (11%).

Table 6 also presents the nonresponse rate as a function of lateral geometry. The significant chi-square test, $\chi^2(2, N = 442) = 33.3, p < .01$, revealed that controller responses were very likely when tracks were laterally converging (93%) but decreased in frequency when tracks were parallel (88%) and when they were diverging (60%; a 40% nonresponse rate).

Three corresponding analyses were then carried out with four categories of vertical geometry (Table 7): nonlevel (N), both parallel (climbing or descending; P), one climbing and one descending so that they were converging (C), or both level (L). There were too few trials with diverging tracks on the vertical axis to include in the analyses. As Table 7 shows, the FA rate was unaffected by vertical geometry, $\chi^2(3, N = 452) = 3.48, p > .10$. Table 7 also shows that the percentage of anticipatory responses was greatly influenced by vertical geometry, $\chi^2(3, N = 342) = 39.8, p < .0001$, with this desirable behavior increasing when geometry was converging (88%) or when both aircraft were level (80%), compared with parallel climbs or descents or nonlevel (mean = 39%). Finally, Table 7 indicates that the response rate did not significantly differ across vertical geometry categories, $\chi^2(3, N = 452) = 4.71, p > .10$. However, it is noteworthy that the same vertically converging category that generated the highest anticipation rate also generated the highest response rate (97% vs. 90% average across the other three categories).

Thus to summarize briefly, automation performance (FA rate) was not influenced by conflict geometry, but controller performance was so influenced.

Table 7: Categorical events as a function of vertical geometry. Measures in proportion (P) of events

Variable	Nonlevel	Parallel	Converging	Level
Vertical geometry, false alert (FA) rate				
P(FA)	0.45	0.52	0.39	0.47
n	194	102	86	70
Vertical geometry, anticipation				
P(anticipation)	0.41	0.37	0.88	0.80
n	178	99	25	40
Vertical geometry, controller response				
P(response)	0.89	0.92	0.97	0.88
n	194	102	86	70

DISCUSSION

The primary purpose of this research was to seek evidence that the relatively high FA rate in the conflict alerting system of ATC en route centers might be responsible for creating a cry wolf syndrome of ignoring or delaying responses to true alerts. We proceeded in four steps: (a) assessing whether in fact there were false alerts in our data base, (b) assessing whether there were many cases of nonresponses associated with an LOS, (c) assessing whether variance in these was associated with variance in cry wolf symptoms, and (d) examining what other aspects of conflicts might be associated with both desirable (anticipatory) and undesirable (LOS) controller behavior and system outcomes.

Our first step, addressed as Hypothesis 1, revealed that indeed there were a large number of alerts that could be categorized as false even when the impact of a subsequent trajectory change was subtracted, a change that we infer was triggered by a controller instruction. We note that this 47% false-alert rate (proportion of alerts that are false) is modest compared with that in some other areas, for example, health care. However, rates are comparable with those involved in weather forecasting (Barnes et al., 2006) or other aspects of ATC alerting (Krois, 1999). There also was substantial variance in this false-alert rate across the five facilities. Although the cause of this variance is not apparent (it does not appear to be related to traffic density or CA frequency), its presence clearly allowed us to test Hypothesis 2, whether centers with a higher FA rate produce more nonresponses to true alerts or delayed responses to all alerts, that is, the predictions of Hypothesis 2b.

With regard to Hypothesis 2b, the evidence was mostly negative. Concerning the most potentially dangerous expression of cry wolf—nonresponse to true alerts—although Figure 4 reveals that controllers at higher-FA centers showed a lower response rate, that lower rate did not produce more LOSes or more nonresponses to true alerts. With regard to response delays, there was no evidence that centers with a higher FA rate showed either later responses or fewer anticipatory responses. Thus, the data in Figure 4 may be interpreted by assuming that in centers where the FA rate is higher, controllers are increasingly ignoring trajectory pairs that are clearly not a danger (even though they did trigger the CA). Ignoring these trajectory pairs seems to reflect an optimal controller strategy.

With regard to Hypothesis 2a, the substantial proportion of alerts that were false (47%) did not appear to be associated with a syndrome in which the controller did not respond to a true alert and produce an LOS. Nonresponses with an LOS constituted only 2/497, or 0.4% of the data, less than the independent product of nonresponse rate and LOS rate would predict. Hence these results provide further evidence that controllers in the current sample were not ignoring the alerts as they might in a cry wolf scenario.

Hypothesis 3 concerns the effects of conflict properties on automation dependence broken down by compliance and reliance (Meyer, 2001, 2004). Compliance is responding to an alert, and so noncompliance is represented by the small population of nonresponses. Reliance describes the tendency to withhold action when the alert is "silent," and hence nonreliance is represented by the large frequency of anticipatory responses (Levinthal & Wickens, 2005).

Given this categorization, we hypothesized (Hypothesis 3a) that factors making conflicts easier to visualize would reduce reliance, because of course, such better visualization would mean that it is easier to process the raw data on the display in parallel with the automation monitoring. This hypothesis was confirmed in two respects. First, trajectory pairs that projected a smaller miss distance (true alerts; more visually evident as a problem) triggered a greater number of anticipatory responses (Table 5). Second, trajectory pairs that were more "visible," or prototypical of conflicts, also triggered more anticipatory behavior. In both the vertical and lateral axes, these prototypical conflict pairs were defined as "converging" conflicts, manifesting negative relative velocity. In the lateral map plane, it is quite evident why this should be the case, given the visibility of the closing tracks on the radar display. In the vertical axis, such convergence is displayed not graphically but instead by the altitude data tags changing in opposing directions.

In terms of reduced compliance (Hypothesis 3b), a false alert–prone system does not appear to have an effect, as described earlier. However, on a trial-by-trial basis, it is evident that a large predicted miss distance (characteristic of a false alert) reduces compliance (fewer responses; Table 2), again strong evidence that controllers are processing the raw data; if it visually appears that a conflict will not take place (large projected miss distance), the CA can be ignored, reflecting noncompliance. However, less readily interpretable is the finding that difficult-to-visualize trials also reduced compliance. In particular, in the lateral axis, diverging conflicts yielded a 40% nonresponse or noncompliance rate (compared to 8% for other conflicts), and in the vertical axis, although there were no CAs triggered by vertically diverging trajectories, those that were converging (most easy to visualize) were associated with only 3% of nonresponses (97% compliance), compared with 10% (90% compliance) for all other categories. Thus, we might assume from those trials that when the conflict does not appear to be visually evident in the raw data, controllers show an increasing tendency to ignore the alerts, although this ignoring does not appear to compromise safety. We consider the practical implications of this interpretation later.

Thus, although the data do suggest ample evidence for false alerts, and that such alerts are less likely to trigger a response (as if they are ignored), it does not appear that these false alerts engender the sort of negative "set" toward the alerting system associated with the cry wolf effect. Four possible explanations for this null conclusion can be offered. First, analyses reported in Wickens et al. (2008) revealed that a majority (80%) of the FAs in the current data were "forgivable" or "acceptable" in that they could be seen as resulting from a slightly conservative threshold of the detection algorithm, that is, erring on the side of more FAs at the expense of minimizing misses or delayed CAs. As such, each occurrence of these acceptable FAs can serve to reinforce the controller's classification of the raw data (for anticipatory responses), resulting in what we might infer to be reinforcing trust in the system (Lees & Lee, 2007).

Second, it is noteworthy that many of the circumstances in which cry wolf behavior is observed have been in distinct dual-task settings—the FA-prone alerting system is supervising a lower-priority background task while the user's attention is heavily focused on a higher-priority task (e.g., Bliss, 2003; Bliss & Dunn, 2000; Dixon et al., 2007; Dixon & Wickens, 2006). In contrast, in the CA system examined here, both the controller and the automation system are addressing the same high-priority task of safe traffic separation management. Therefore, task switching is not required between the high-priority CA detection and something else. As a consequence, any loss of compliance that may result does not (and cannot) lead to the sort of task-switching delay, observed by Dixon and Wickens (2006; see also Wickens, Dixon, Goh, & Hammer, 2005), involved when the pilot is flying the aircraft (primary task) while monitoring for traffic (secondary task). This is because such task switching in the current ATC context simply does not take place when the primary task is also the alerting task, nor does it lead to the sort of concurrent task interference that will increase controller frustration with the alerting system.

Third, we note that the CA system here was associated not with sound—sound alerts being more intrusive and hence more potentially annoying when false—but rather with the flashing displayed data blocks.

Fourth, it is certainly possible that such an effect may have been present but was offset by other factors that varied in a confounding fashion between the centers. Although it does not appear that workload (traffic load) was such a factor (see earlier discussion), other possibilities will be discussed later. Finally, it should be noted that one influence that can sometimes be invoked to account for finding null effects, low statistical power, can probably be ruled out here because of the high-N study, including almost 500 observations.

Although the data do reveal a 10% LOS rate, it is important to note that these are not automatically considered to be "operational errors" that trigger some form of sanctions on the responsible controller (although we do not have access to data on the linkage between these LOSs and operational errors). A formal LOS as defined by MMR < 1.0 may trigger further scrutiny by an FAA supervisor, but such scrutiny can often reveal no sense of imminent danger, particularly if the MMR value is close to 1.0 (which it was for most of the LOS events in our sample; Wickens et al., 2008).

Practical implications and limitations. The main practical implications of this study are twofold. First, it appears that this particular system (conflict prediction in the en route airspace) does not warrant any substantial modifications to address its false-alert rate. The rate, although modestly high, appears not to contribute to any deficit in controller or system performance. Second, there is some evidence that certain geometries are more difficult for the controller to visually process than others (those with less convergence and more divergence on either axis) and that such difficulty may reduce compliance. In developing future systems, it might be worthwhile to consider reducing the CA threshold and perhaps amplify its salience when such geometries appear, thereby rendering it more likely that the alert information is complied with.

There were several limitations and constraints in this study. First, we caution generalizing the current results and conclusions—drawn from en route ATC centers—to CAs in terminal airspace, where FA rate may be considerably higher (FAA, 2006) and the boundaries between LOS and non-LOS are far less clear-cut. Similar caution should be exercised in generalizing to terrain alerts, also the target of the original NTSB (2006) report.

Second, we note the obvious fact that because this was not a true experimental study and we did not have direct access to participants (e.g., to assess controller trust or to examine their direct responses, or voice communications), many aspects of our interpretation involve weaker inferences than could have been made with a true experimental study. Hence we must acknowledge the lack of certainty of our conclusions. Despite this fact, at least one potential and serious confound in the study can be partially addressed: Centers with a higher false-alert rate might be those with a higher workload (e.g., greater traffic density), and this, rather than the FA rate, could be the source of effects. We address this by noting first that we did not observe the expected cry wolf effect. (Had we done so, it would have been more necessary to demonstrate that increasing workload was not responsible.) Furthermore, we actually found evidence that the busiest center produced less cry wolf behavior, as inferred by a shorter RT.

A third limitation is that the data provided to us were from only the heaviest workload period within each center. Therefore, it is possible that automation dependence might have been amplified in the data sampled, relative to what it would have been during other periods. This is because other analyses (Wickens & Dixon, 2007) indicate that high workload appears to increase dependence on automation. In this light, it is all the more interesting that a high level of dependence, reflected in cry wolf behavior, was not observed.

Fourth, we consider two other confounds that might have offset and therefore masked an existing cry wolf influence: cultural differences between centers and between levels of experience. With regard to culture, there is no doubt that centers, like people, are not homogenous. Some may foster greater belief in, and dependence on, the advice of automation (Wickens, Mavor, & McGee, 1997). Therefore, the pattern of data could be accounted for by assuming that centers with a higher FA rate were more accepting (trusting) of the CA system to offset the possible inhibiting effect of the higher FA rate. We have no good way of assessing this possibility.

With regard to training and experience, many researchers of the cry wolf effect regard it as a state of losing trust that develops over time, with repeated exposure to FAs. Although we could not assess differences in controller experience level between centers, we simply assume that it is fairly homogeneous and that the development of mistrust would take place across many fewer trials than the amount of experience of the average controller. If anything, more experienced controllers would tend to occupy the busier centers, creating an effect that would amplify a cry wolf effect rather than offset it.

Notwithstanding the potential shortcomings and limitations of this naturalistic study acknowledged previously, we believe that the benefits produced by its large *N* and basis on fully "live data" from controllers with real-world motivations and expectancies offset those shortcomings. Ultimately, the results make a compelling case that more controlled simulation experiments be carried out in the future to confirm these findings.

And what of the cry wolf effect? In spite of the generally negative evidence here, we fully believe that it is present elsewhere, both within and beyond the laboratory (Barnes et al., 2006; Bliss, 2003; Seagull & Sanderson, 2001; Xiao et al., 2004). The challenge of researchers now is to explore the circumstances (such as single- vs. dual-task situations described earlier or "good" vs. "bad" false alerts examined by Lees & Lee, 2007) that can modulate its influence.

ACKNOWLEDGMENTS

This research was supported by Grant No. OGC #20070271 from the Federal Aviation Administration (FAA) to New Mexico State University (NMSU). We acknowledge support from the FAA of Kenneth Allendoerfer, who provided the radar data and gave us advice on many aspects of the analysis, and from Dino Piccioni, the scientific-technical monitor. We acknowledge the contributions of Ken Leiden, Jill Kamienski, Tim Bagnall, and Angie Sebok from Alion Science for their contributions to many aspects of the analysis. We acknowledge the assistance of Amy Wells from NMSU. We also gratefully acknowledge the contributions of two anonymous reviewers.

REFERENCES

Ahlstrom, V., & Panjwani, G. (2003). *Auditory alarms in airways facilities environment* (DOT/FAA/CT-TN04/04). Atlantic City Airport, NJ: Federal Aviation Administration, William Hughes Technical Center.

Allendoerfer, K., & Friedman-Berg, F. J. (2007). *Human factors analysis of safety alerts in air traffic control* (Final Report DOT/FAA 07/22). Washington, DC: Federal Aviation Agency.

Barnes, L. R., Gunfest, E., Hayden, M. H., Schultz, D. M., & Benight, C. (2006). False alarms and close calls: A conceptual model of warning accuracy. *Weather and Forecasting, 22*, 1140–1147.

Bliss, J. (2003). An investigation of alarm related accidents and incidents in aviation. *International Journal of Aviation Psychology, 13*, 249–268.

Bliss, J., & Dunn, M. (2000). The behavioral implications of alarm mistrust as a function of task workload. *Ergonomics, 43*, 1283–1300.

Breznitz, S. (1983). *Cry-wolf: The psychology of false alarms.* Hillsdale, NJ: Lawrence Erlbaum.

Burns, C. (2006) Proactive monitoring in the petrochemical industry. *Safety Science , 44*, 27–36.

Dixon, S. R., & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle flight control: A reliance-compliance model of automation dependence in high workload. *Human Factors, 48*, 474–486.

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse the misses? *Human Factors, 49*, 564–572.

Dixon, S. R., Wickens, C. D., & Seppelt, B. (2005). Auditory preemption versus multiple resources: Who wins in interruption management? In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 463–467). Santa Monica, CA: Human Factors and Ergonomics Society.

Federal Aviation Administration. (2003). *Common ARTS computer program functional specification for conflict alerts* (NAS-MD-632). Washington, DC: Author.

Federal Aviation Administration. (2006). *Human Factors Study of Air Traffic Control Safety Alerts white paper and progress report.* Washington, DC: Author (2007).

Friedman-Berg, F., Allendoerfer, K., & Pai, S. (2008). Nuisance alerts in operational ATC environments: Classification and frequency. In *Proceedings of the 52nd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 104–108). Santa Monica, CA: Human Factors and Ergonomics Society.

Getty, D., Swets, J., Pickett, R., & Gonthier, D. (1995). System operator response to warnings of danger. *Journal of Experimental Psychology: Applied, 1*, 19–33.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, UK: Wiley.

Krois, P. (1999). *Alerting systems and how to address the lack of base rate information*. Unpublished manuscript, Federal Aviation Administration, Washington, DC.

Kuchar, J., & Yang, L. C. (2000). A review of conflict detection and resolution modeling methods. *IEEE Transactions on Intelligent Transportation Systems, 1*,179–189.

Lawless, S. T. (1994). Crying wolf: False alarms in a pediatric intensive care unit. *Critical Care Medicine, 22*, 981–985.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*, 50–80.

Lees, N., & Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics, 30*, 1264–1286.

Levinthal, B., & Wickens, C. D. (2005). *Supervising two versus four UAVs with imperfect automation: A simulation experiment* (AHFD-05-24/MAAD-05-7). Savoy: University of Illinois, Aviation Human Factors Division.

Madhavan, P; Wiegmann,m D. A., & Lacson, F. C. (2008). Automation failures on tasks easily performed by operators undermine trust in automated aids. *Human Factors, 48*, 241-256.

Maltz, M., & Shinar, D. (2003). New alternative methods in analyzing human behavior in cued target acquisition. *Human Factors, 45*, 281–295.

Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors, 43*, 563–572.

Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors, 46*, 196–204.

Meyer, J., & Bitan, Y. (2002). Why better operators receive worse warnings. *Human Factors, 44*, 343–353.

NTSB (2006). National Transportation Safety Board Safety Recommendation A-06-44 through A-06-47. Washington DC: NTSB.

Paglione, M., Ryan, H., & Liu, S. (2007). *Evaluation of en route host computer system's tactical alert processing* (DOT/FAA/TC-TN07/13/). Washington, DC: Federal Aviation Administration.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors, 29*, 695–706.

Parasuraman, R., & Riley, V. A. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*, 230-253.

Rice, S. (in press). Examining single and multiple-process theories of trust in automation. *Journal of General Psychology*.

Seagull, F. J., & Sanderson, P. M. (2001). Anesthesia alarms in context: An observational study. *Human Factors, 43*, 66–78.

Sorkin, R. D. (1989). Why are people turning off our alarms? *Human Factors Bulletin, 32*(4), 3–4.

Stanton, N., & Babar, C. (1995). Alarm-initiated activity. *Ergonomics, 38*, 2414–2431.

Swets, J. A., & Pickett, R. M. (1982). *The evaluation of diagnostic systems*. New York: Academic Press.

Thomas, L. C., & Wickens, C. D. (2008). Display dimensionality and conflict geometry effects on maneuver preferences for resolving in-flight conflicts. *Human Factors, 50*, 576–588.

Vashitz, G., Meyer, J., Parmet, Y., Peleg, R., Goldfarb, D., Porath, A., *et al.* (2008). Defining and measuring physicians' responses to clinical reminders. *Journal of Biomedical Informatics, 42*, 317–326.

Wickens, C. D., & Dixon, S. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomic Science, 8*, 201–212.

Wickens, C. D., Dixon, S. R., Goh, J., & Hammer, B. (2005). Pilot dependence on imperfect diagnostic automation in simulated UAV flights: An attentional visual scanning analysis. In *Proceedings of the 13th Annual International Symposium of Aviation Psychology* [CD-ROM]. Dayton, OH: Wright State University.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Mavor, A. S., & McGee, J. P. (Eds.). (1997). *Flight to the future: Human factors in air traffic control*. Washington, DC: National Academy Press.

Wickens, C. D., Mavor, A. S., Parasuraman, R., & McGee, J. P. (Eds.). (1998). *The future of air traffic control: Human operators and automation*. Washington, DC: National Academy Press.

Wickens, C. D., Rice, S., Keller, M. D., Hughes, J., Hutchins, S., & Clayton, K. (2008). *Addressing the alert problem in ATC Facilities: Final report*. Las Cruces: New Mexico State University.

Wickens, C. D., Levinthal, B. & Rice, S. R. (in press). Imperfect reliability in unmanned air vehicle supervision and control. In A. W. Evans (Ed.), *Human-robotics interaction in future military systems*. Brookfield, VT: Ashgate.

Woods, D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics, 38*, 2371–2393.

Xiao, Y., Seagull, F. J., Nieves-Khouw, F., Barczak, N., & Perkins, S. (2004). Organizational-historical analysis of the "failure to respond to alarm" problems. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 34*, 772–776.

Xu, X., Wickens, C. D., & Rantanen, E. M. (2007). Effects of conflict alerting system reliability and task difficulty on pilots' conflict detection with cockpit display of traffic information. *Ergonomics, 50*, 112–130.

Christopher D. Wickens is a senior scientist at Alion Science Corporation, Micro Analysis and Design Operations, in Boulder, Colorado, and professor emeritus at the University of Illinois at Urbana-Champaign. He received his PhD in psychology from the University of Michigan in 1974.

Stephen Rice is an assistant professor of psychology at New Mexico State University in Las Cruces. He received his PhD in engineering psychology from the University of Illinois at Urbana-Champaign in 2006.

David Keller is a postdoc in the Engineering Psychology Department at New Mexico State University in Las Cruces, where he received his PhD in engineering psychology in 2009.

Shaun Hutchins is a senior human factors engineer at Alion Science, Micro Analysis and Design

Operation. He received his MA in experimental psychology from New Mexico State University in Las Cruces in 2007 and is currently a PhD student studying research methodology at Colorado State University.

Jamie Hughes is a PhD candidate in the Social Psychology Department at New Mexico State University in Las Cruces. She received her MA in social psychology at Illinois State University in Bloomington in 2006.

Krisstal Clayton is a PhD candidate in the Social Psychology Department at New Mexico State University in Las Cruces, where she received her MA in social psychology in 2006.

Date received: January 30, 2009
Date accepted: July 2, 2009