



## On the theory of fuzzy signal detection: theoretical and practical considerations

PETER A. HANCOCK†\*, ANTHONY J. MASALONIS‡ and RAJA PARASURAMAN§

† Department of Psychology, University of Central Florida, Orlando, Florida 32186, USA

‡ Center for Advanced Aviation System Development, the MITRE Corporation, McLean, VA 22102, USA

§ Cognitive Science Laboratory, The Catholic University of America, Washington, DC 20064, USA

*Keywords:* Signal detection theory; Fuzzy set theory; Human performance; Machine performance.

This work examines the foundations for and explores the implications of fuzzy signal detection (Fuzzy SDT), a theory that represents the marriage of two powerful extant theories, fuzzy set theory and signal detection theory. Fuzzy SDT permits the modelling and prediction of human, machine, and human-machine performance in a wide range of settings. Fuzzy SDT exploits the strengths of each theory to provide new and dynamic insights into performance. Fuzzy SDT explicitly recognizes that the binary decision states of classic signal detection represent two ends of a single continuum whose uncertainty decreases towards such end states and is maximized in its centre. It is shown how Fuzzy SDT has its origins in some more general concepts of human performance, and companion works are referenced which provide the mathematical foundation for Fuzzy SDT and its application in a specific domain. The present work examines the wider implications of Fuzzy SDT by illustrating the relevance of fuzzification in the larger cycle of design, configuration, and use of technology. It also examines the broader concerns of the temporal relationship between signal and response, showing time to be a crucial, if neglected, dimension of action, the exploration and exploitation of which can produce a deeper understanding of human behaviour in psychology, machine behaviour in engineering and human-machine behaviour in ergonomics.

### 1. Introduction

Whilst one might desire that the world be presented in terms which are certain, it is ubiquitously the case that observers have to recognize and cope with ambiguity. Such ambiguities are spread over space and time and are often resolved by an individual's active exploration. When phenomena exceed one's simple, unaided exploratory capabilities, they provide a strong stimulus for technological innovation. As Hancock (1997) has noted, 'while the perception-action link may explain *how* we explore our environment, the perception-action gap may explain *why* we explore in the first place'. In typical terrestrial environments, stimuli resolve over space and time such that an observer's understanding of what is present and what response is appropriate evolves dynamically. At some juncture, observers reach a sufficient degree of certainty concerning the state of the world around them and act upon

---

\* Author for correspondence. e-mail: phancock@pegasus.cc.ucf.edu

that decision. The proposition that action can control perception is a result of this view in which behaviour is continual and situated in the stream of reality (Powers 1973, Flach *et al.* 1995).

This capacity, to reduce uncertainty through exploration, is reflected in the assertion that an object or concept and its opposite cannot co-exist. The corollary, that the world can be divided into mutually exclusive categories, has been vital in human evolution. It is indeed a fundamental characteristic of most human achievements. For example, the very foundation of mathematics is predicated upon the concept of number, and number is only possible if concrete items or objects are seen as sufficiently alike so that they may be grouped together in a set and others, by definition, excluded from that set. Thus, if there were no primitive perceptual concept of set, the whole basis for arithmetic itself would be compromised. The development of number and, much later, formal set theory itself, is, thus, contingent upon intrinsic perceptual capabilities. However, perception itself is a complex and non-stationary ability with significant variability between individuals. What one observer classes as 'signal', another observer will see as 'noise'. Each may wish to separate the wheat from the chaff; but they do so in very different ways. Therefore, what constitutes the 'world' is always subject to the filter of perception, and so the nominal 'real-world' may indeed appear different from individual to individual (Hume 1739).

Tanner, Swets, Green and their colleagues formalized these properties of the detection of signals and noise and developed a mathematical approach now universally known as signal detection theory (SDT: Tanner and Swets 1954, Green and Swets 1966). SDT has been widely applied in the evaluation of the accuracy of diagnostic systems that seek to distinguish signal from noise (see for example Swets and Pickett 1982, Sorkin and Woods 1985, Swets 1977, 1996). Due to the range and success of SDT, this approach has been rightly characterized as one of the most robust theories in the behavioural sciences and today it remains one of the most powerful theoretical and also practical constructs in all of the human performance literature. SDT has been used in a wide variety of settings, even as an investigative tool to evaluate claims of paranormal activity (Jensen 1989). SDT divides the observable world into two components. The first, designated the 'signal' represents a state of the world in which the pre-agreed target item or event is actually present as a true 'state of the world'. The second categorization, 'noise' represents background fluctuation of the environment which is, thus, composed of any form of stimulation which is not the agreed 'signal'. Of course, what constitutes a 'signal' is largely arbitrary and is the subject of a pre-agreement between parties involved in the detection process. As there is always some form of extraneous stimulation present, the real distinction is between 'noise' and 'signal plus noise'. Such a display, or repeated displays are given to an observer who is forced into a choice between either 'yes' the signal is present or 'no' the signal is not present (see Macmillan and Creelman 1991). When the observer's response accords to the true state of the world, it is classed as a 'hit' when the signal is present and a 'correct rejection' when there is only noise. However, observers are not always right. When the difference between the signal and the background is small and particularly when observers are stressed, tired, bored, or frustrated or when many repeated responses are required, observers make mistakes (see Davies and Parasuraman 1982, Hancock 1984, Warm 1984, Hancock and Warm 1989). When the observer responds positively but there is in reality no signal present, one has a 'false alarm', and when the observer responds negatively but a signal is actually present one has a 'miss'. Through examining the

combination of these four outcomes over repeated trials, one can provide an estimate of how good the observer is in terms of quantitative measures of sensitivity and response bias. The observer might be a human but equally well could be a machine, since observation as a process is not restricted to living things. As a result of the quantitative analysis that this procedure provides, it has proved particularly appealing to those with an engineering perspective, since 'real' numbers provide crucial input into formal models and, thus, considerable comfort to those who deal with the world in that fashion.

## 2. Some limits of signal detection theory

Despite the tremendous and deserved success that SDT has garnered, there are a number of factors which restrict its application in practical terms, and these derive from some intrinsic theoretical limits. Two such limitations are discussed below.

### 2.1. *Knowing the state of the world*

Traditional SDT requires the mapping of environmental events or sources of evidence, the 'evidence variable' into two categorical states of the world that do not overlap. In the laboratory, this is accomplished by the researcher, who controls the stimuli which compose the experimental 'world'. However, in the real-world, such a mapping is fuzzy rather than discrete (see also Karwowski 1992). In most real-world settings, the definition of a signal is context-dependent and varies with a plethora of factors. Thus, to apply SDT in the real-world, one essentially has to know what the true state of the world is. However, if this state was known *a priori*, then actual detection would be obviated. Conversely, to know the true state of the world one has to engage in some form of detection in the first place. Whilst it is true that more and more one is 'creating' the technical environments in which one works, this 'circularity' of signal definition limits SDT. In fairness, of course, the question of the true state of the world will plague any proposed detection theory. However, the problem of non-stable categorization of signal and noise is fundamentally a temporal one, and it is the dimension of time which represents the major concern.

### 2.2. *The temporal dimension of signal detection*

In the authors' view, the greatest limitation to the ubiquity of SDT is the problem of time. As the fundamental dimension of existence, time can not be halted or rolled back. Events are conceived as occurring 'in time', or time is viewed as an emergent property of event occurrence (cf., Gibson 1975, Hancock 1993). Thus, *when* a decision or detection is made is at least as important as *what* decision or detection is made. At present, current formulations of SDT in the main do not address the temporal dimension directly. For example, in classic SDT, there is no specific mechanism to determine the latency of response. Whilst hit, miss, false alarm, and correct rejection are the four possible outcomes, nothing indicates *when* these responses will occur, either within an individual response or across a series of responses. Typically, this is because the methods used in SDT research artificially influence the time factor, either by temporal limitations on response or by some other experimental manipulation which quizzes observers in the spatial domain but restricts them in the temporal domain. In the real world, there may indeed be time limits imposed by the natural constraints of a task at hand. Often, as in competition chess, a time limit has to be imposed, since endless prevarication renders the very phenomenon moot. However, when to respond is an observer choice not an

experimenter's decision. Thus, in SDT, when the detection response occurs is unspecified and so represents a crucial missing piece of the jigsaw in behavioural prediction. SDT, therefore, takes a photograph or 'snapshot' of the signal/response event. Further, SDT is typically applied when an individual or a machine engages in a series of observations or judgements, giving repeated opportunities for evaluating their capability. In such circumstances, SDT provides much information *on average*, about the summed ability to respond across trials. What is missing is the prediction of the sequential effects embedded in a series of responses. That is, where do hit, miss, false alarm, and correct rejection occur sequentially in the response stream as differing states of the world are presented for consideration. The purpose here is to address such questions and to provide a construct which may help in their resolution.

### 2.3. Purpose of this paper

The fundamentals of Fuzzy Signal Detection Theory have been described in Parasuraman *et al.* (2000). This paper elaborates on this hybrid theory to address the fundamentals of detection response. Through the combination of fuzzy set theory (Zadeh 1965) and classical SDT, an approach is posited to provide avenues through which the calculation and prediction of detection behaviour can be evaluated in real settings. Extensions of Fuzzy SDT are also considered using other theoretical approaches such as catastrophe theory (see Thom 1975, Zeeman 1977) to discuss broader implications to a general formulation on the prediction of human behaviour.

### 3. Fuzzy signal detection theory: basic postulates

Whilst SDT has proved an important analytical technique, fuzzy set theory has likewise represented an important quantitative approach to capture the understanding of phenomena. Although not without its critics (Gardner 1995), fuzzy set theory has proved helpful in many realms (see Kosko 1993, 1997), including issues related to ergonomic concerns (see Chignell and Hancock 1986, Karwowski and Mital 1986). Given the value of these two approaches, it may be anticipated that a union of the two techniques would represent a valuable theory and analytic methodology. Parasuraman *et al.* (2000) have proposed a direct application of fuzzy logic to the classical methods of SDT and provided examples of how the traditional SDT parameters can be calculated when the degree to which a signal has occurred and the degree to which a response has been made are not collapsed or 'rounded' to 0 or 1 at any point in the analysis. The fuzzy SDT analysis developed by Parasuraman *et al.* (2000) involves four steps, which are outlined below:

- (1) *Selection of mapping functions for states of the world (SOW) and responses.* Whereas, in traditional SDT, the SOW is divided in a binary fashion into either signal or noise, Fuzzy SDT assigns each possible SOW to a degree of signal,  $s$ , between 0 and 1. Thus, a particular set of physical variables corresponding to the SOW may represent a signal, but only to a degree, say  $s = 0.7$ . If the SOW is less signal-like and more noise like,  $s$  may have a smaller value, say  $s = 0.2$ . The function that maps the SOW to  $s$  is known as the *mapping function*. Each possible response that can be made by the observer to events can also be assigned a response degree  $r$  between 0–1. The response degree might correspond to the observer's confidence that a signal occurred, or to some measure of reported signal criticality or intensity. For

example,  $r = 0.9$  could represent a high degree of confidence that the event was a signal.

- (2) *Use of implication functions.* In traditional SDT, logical IF-THEN functions can be used to determine the outcome categories of the detection system. For example, if the event is a signal ( $s = 1$ ) and the observer's response is positive ( $r = 1$ ), then the outcome is a 'Hit'. That is, the outcome category of hit has a value of  $H = 1$  and the remaining three categories, misses ( $M$ ), false alarms ( $FA$ ), and correct rejections ( $CR$ ) have values of 0. Similarly, if the observer responded negatively ( $r = 0$ ), then  $M = 1$ , and  $H$ ,  $FA$ , and  $CR$  are all 0. In Fuzzy SDT, however, a given event-response pair may belong to some degree to *more than one* outcome. For example, suppose  $s = 0.9$ , i.e. the SOW strongly but not absolutely points to a signal. Suppose the observer responds with  $r = 0.8$ , i.e. strongly but not unequivocally responds positively that the event was a signal. Then, whereas traditional SDT would classify the outcome as only a hit ( $H = 1$ ), Fuzzy SDT classifies the outcome as mostly a hit, but also, to a small degree, a miss and a correct rejection. Parasuraman *et al.* (2000) proposed a set of implication functions for deriving the fuzzy set memberships of the four outcomes, as follows:

$$\begin{array}{ll} \text{Hit:} & H = \min(s, r) \\ \text{Miss:} & M = \max(s-r, 0) \\ \text{False alarm:} & FA = \max(r-s, 0) \\ \text{Correct rejection:} & CR = \min(1-s, 1-r) \end{array}$$

Thus, in this example, whereas in traditional SDT,  $H = 1$  and  $M = FA = CR = 0$ , in Fuzzy SDT,  $H = 0.8$ ,  $M = 0.1$ ,  $FA = 0$ , and  $CR = 0.1$ . Note that, in both traditional and fuzzy SDT, the outcome memberships always sum to 1. In essence, Fuzzy SDT partitions some of the categorical set membership of traditional SDT ( $H = 1$ ) into the other outcomes.

- (3) *Computation of fuzzy hit and false alarm rates.* This step involves the calculation, across  $n$  trials or observations, of the weighted average of the set membership values for two of the four outcomes: Hits and False alarms. The weight for Hits is the average of the signal membership values ( $s$ ) across the  $n$  trials; that for false alarms is the average of noise membership values ( $1-s$ ) across the  $n$  trials.
- (4) *Computation of fuzzy sensitivity and criterion measures.* This step is straightforward and involves the computation of sensitivity (e.g.  $d'$ ) and criterion (e.g.  $\beta$ ) measures from the fuzzy hit and false alarm rates computed in step 3. This step is essentially the same as in traditional SDT.

Using a reanalysis of data from two studies involving aircraft conflict detection in air traffic control (ATC), Masalonis and Parasuraman (2000) have demonstrated that Fuzzy SDT provided additional insight in detection performance over that derived from traditional SDT. Analysis of detection performance of an automated conflict detection system in a low-signal, real-world environment resulted in a lower value of hit and false alarm as compared with the traditional SDT method. Fuzzy SDT resulted in a lower hit rate and false alarm rate because it more accurately captured the fact that a number of 'signals' (near-conflicts between aircraft) were present. Such events may have required attention, despite not being full-fledged

signals according to an arbitrary cutoff. Masalonis and Parasuraman acknowledge the possibility that their results were a function of the way that fuzzy values of  $s$  (signal) and  $r$  (response) were assigned for the analysis, and that different assignments might have resulted in different findings when comparing Fuzzy SDT and traditional SDT. Most important, the Fuzzy SDT analysis provided points to potential areas of improvement in the performance of the automation not apparent from traditional SDT analysis.

#### 4. Time and fuzzy SDT

##### 4.1. *The temporal component*

Parasuraman *et al.* (2000) sought to establish the benefits of allowing the ‘gray areas’ of signal and response definition into SDT analyses. In addition to the gradations that are possible in the objective definitions of these parameters, an additional fuzziness is injected by probability (Karwowski 1992). Whatever the extent of the fuzziness of the *actual* state of the world, the *perception* of that state can also be fuzzy. This is probably more true of a detector running on analogue hardware (such as a human), but can, in principal, be true of any detector. One way this component of fuzziness can be represented is as a *probability* that the stimulus is a signal. This probability judgement may change over time as more information is acquired. Stimuli, and/or perceptions thereof, resolve over space and time such that an observer’s understanding of what is being presented and what response is appropriate, evolves dynamically. Under these conditions, the ‘signalness’ of the environment and the response that ought to be emitted remain uncertain and change as information is assimilated. Traditional SDT can not deal directly with these dynamic cases, being essentially an *atemporal* method that represents a momentary ‘snapshot’ judgement. The usual finesse of this problem is to repeat a SDT analysis many times. This may be possible in the experimental laboratory but is much more problematic in real-world situations. This paper discusses the temporal component of SDT by allowing that the mathematical characteristics of sensation and perception of a given stimulus can change over time as information is accumulated. In essence, as has been noted, the momentary judgement taking place in traditional SDT takes a black and white ‘snapshot’ of events. Fuzzy SDT provides a photograph in vibrant colour but is, at this stage, still a static representation. Adding the temporal component captures the dynamics inherent in most real-world decision-making situations. Temporal Fuzzy SDT takes a ‘movie’. In the real-world, events unfold in time. Thus, the value of a signal ( $s$ ) varies between high and low values but rarely reaches an unequivocal value of either 0 or 1. Thus, as has been noted (Parasuraman *et al.* 2000), a more complete description of the signal mapping function based on SW (state of the world) would not be  $s=f(\text{SW})$ , but rather  $s=f(\text{SW}, t)$ , although in reality  $t$  is always a world property.

##### 4.2. *The timing of response*

A tradition in experimental psychology publications, not ubiquitous but very common, is to divide results into two sections, one reporting the effects of the independent variables of interest on response time for those trials that resulted in a hit, and another, often briefer, missive reporting the number and/or percentage of misses, false alarms and correct rejections. In a task with time urgency, which to some degree describes the vast majority of laboratory and real-world tasks, speed-accuracy trade-offs come into play. In the laboratory, one can *constrain* a time limit for judgement,

but this is a *paradigmatic* manipulation, not an attribute of SDT *per se*. What is a miss? Typically, a miss is considered as a lack of response *after an arbitrary period of time*. What if the observer were 'simply thinking about it'? Framed in the perspective of a real-world problem, there is usually a time limit associated with a decision being made (e.g. in driving a car it is far preferable to decide upon, execute, and complete the braking manoeuvre *before* the vehicle impacts the brick wall). However, faster is not always necessarily better, especially if there are costs associated with a false alarm (brake wear and tear, embarrassment, or, most significantly, being rear-ended). These factors are also important when detection is assisted with a collision-avoidance system that has its own detection sensitivity and latency (Parasuraman *et al.* 1997). In the real world, the detection system, because (s)he/it/they are trying to accumulate the most accurate possible assessment of the environment before deciding, may not respond until it is absolutely necessary. Indeed, the same may occur in a laboratory task, depending on the degree to which speed and accuracy are prioritized. Under these circumstances, the difference between a miss and a 'delayed hit' might be difficult to distinguish unambiguously.

One of the central questions that should be resolved in detection behaviour is when a response will occur. In using Fuzzy SDT, one has the opportunity to explore this question in depth. Hancock and Pierce (1989) indicated that there is indeed such a dependency and that this relation could be demonstrated through a linkage between SDT and catastrophe theory (CT) (see Cobb 1981, Stewart and Peregoy 1983, Guastello 1984, Guastello and McGee 1987; but also see Sussmann and Zahler 1978), where repeated *CRs* can lead to *FAs*, due to the expectancy (following numerous non-signal events) that it (the signal) 'has to happen sometime'. To the authors' knowledge, these within-response and across-response temporal dependencies have not been explored to date. It is proposed here that, for an individual response, the timing of that response is dictated by the ratio of the Fuzzy-set membership functions in the major categories of signal detection. Since it is easier to illustrate this in the case of making a response, i.e. a hit or *FA*, rather than inhibiting a response, i.e. a miss or *CR*, the former cases will be dealt with here. Imagine that a display has been presented in which the discriminative difference between noise and signal plus noise is at or near detection threshold. Obviously, the closer to threshold, in general, the longer the time taken to respond, where such temporal freedom is permitted. It is suggested that evidentiary accumulation occurs until the ratio of the hit vs. *FA* membership function fractures some preset value (referring to the case of experience and previous responses). At that point in time, the response is made which represents that persuasion. In terms of catastrophe theory, or more recently complexity theory, the response 'falls' into one of the attractor wells. Clearly, to fully articulate this notion, one would have to specify where the individual 'started' on the attractor space and what the relevant ratio for response would have to be. Further, this would have to be a very strong *a priori* specification since a model with the potential degrees of freedom this contains would provide seductive *post hoc* interpretational appeal. However, in principle, this prior specification could be achieved.

In addition to knowing the momentary value for the response ratio, one would have to understand the nature of the 'surface' of response upon which the individual is operating. Therefore, a sequential dependency in the stream of responses is postulated, or, more generally, a detection system's propensity to respond is predicated upon previous experience. Of course, with clearly deterministic machine systems that do not alter such values, no learning is possible. Such systems are static and, whilst

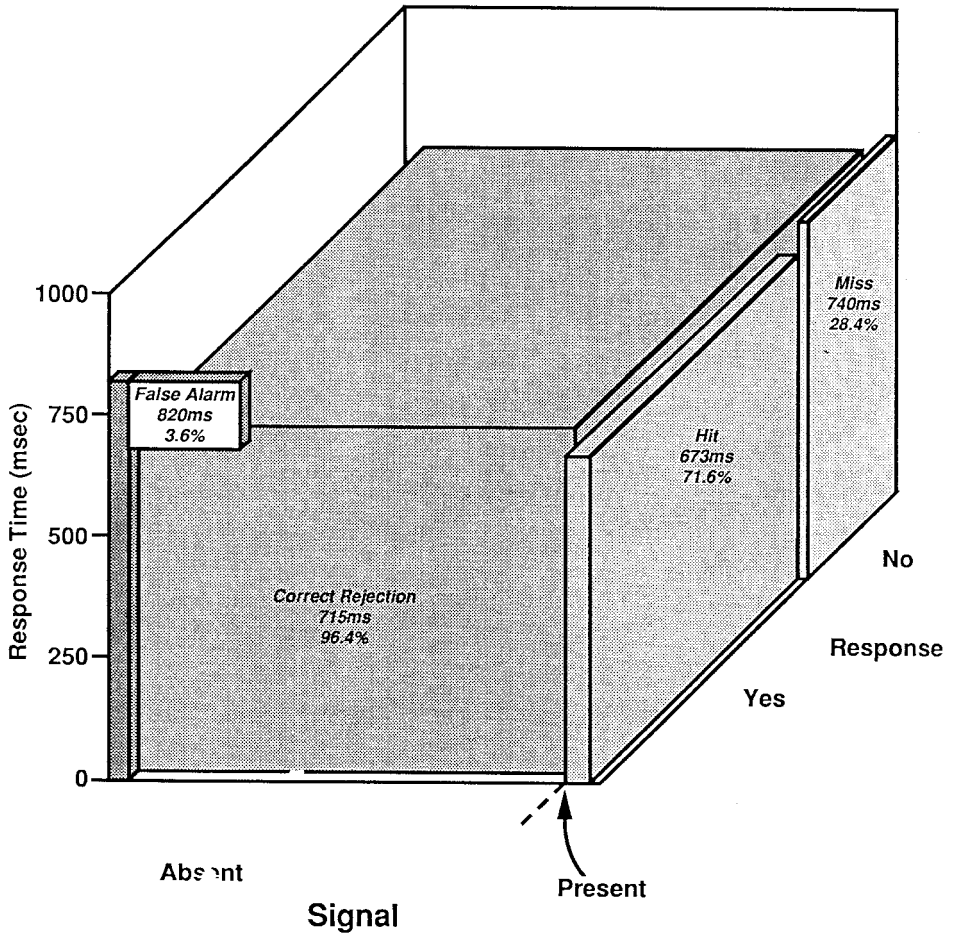


Figure 1. Response surface of temporal latency of hits, misses, correct rejections and false alarms. Redrawn from Parasuraman and Davies (1976).

they do not vary in quality of detection, they also have no manner of improving. Given the case of a system (like a human being) that learns, is there any data which might provide such insights into what the 'surface' of response looks like? Fortunately, there is. Parasuraman and Davies (1976) conducted a vigilance experiment in which classic signal detection was employed that derived response times in all four categories. As with most sustained attention tasks, the actual signal rate was very low and, thus, the number of non-signal events was high and subjects had predominantly to respond with a rejection. Whilst this is true as a general character of vigilance, it need not necessarily be true of all detection tasks and, indeed, the response 'surface' generated can, in some sense, be said to characterize the type of task under consideration. The response surface as derived from these data is shown in figures 1 and 2.

It is proposed that, crude as these surfaces are, they show how the sequential response between the four categories might be related. The surfaces are necessarily 'crude', since there are only four possible categories of response in classic SDT. However, it is postulated that when Fuzzy SDT is applied to this notion of surface

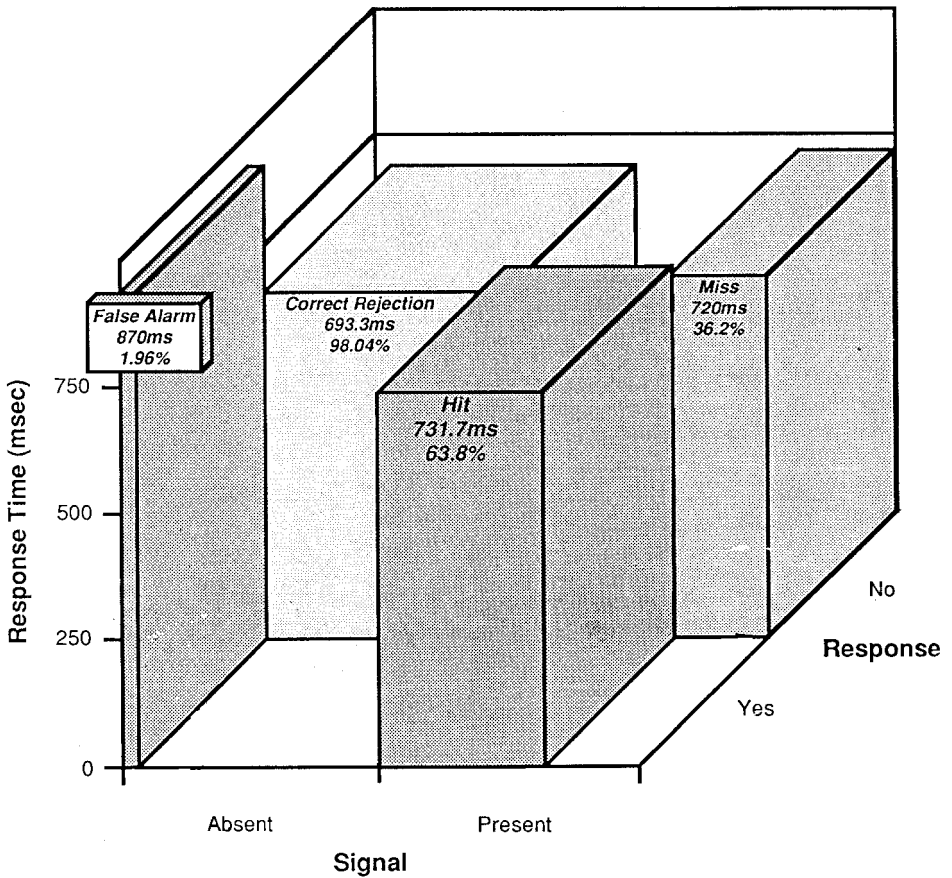


Figure 2. Response surface of temporal latency of hits, misses, correct rejections and false alarms. Redrawn from Parasuraman and Davies (1976).

topology then the 0/1 nature of the collapsed response is obviated and the full complexity of the response surface may be revealed. Of course, after a deterministic response occurs, the collapse to either a yes or no response state is inevitable. From figure 1, it is suggested that, in this typical vigilance environment, the stable state is correct rejection. Thus, for a sequence of responses, contingent upon the true 'state of the world' as dictated by the experimenter (or a random number table), a prolonged series of response inhibitions are required. However, these inhibitions are not neutral events. In this proposal, these sequential inhibitions act to change the threshold of the ratio between fuzzy categories so that, in the present case, a positive response becomes ever more likely. At some point along the sequence, one of two events occurs. Either an actual target comes up, in which case the ratio for the positive response is triggered (being a hit), or a display is presented in which an apparent sufficiency of evidence appears to occur, so that a false alarm is triggered. Note also that such a triggering is crucially dependent upon experience. Thus, in an experimental setting the occurrence of a false alarm is based on how 'trigger happy' the participant is—more formally their  $\beta$ -level.  $\beta$  is directly dependent upon the training of what constitutes a signal early in the procedure and the history of pro-

gression (and of course feedback) later in the procedure which will calibrate the participant's expectation about the overall base rate of a signal and about the likelihood that the next event will contain a signal. All the same factors, of course, are at work in real-world settings.

At fast times, almost any response that is made will be a false alarm—most likely, a false button press. If many of the responses at very fast RTs (times faster than a minimum simple RT of, say, 150 ms) were Hits, one might assume the subject was guessing as opposed to actually detecting, processing, and responding to the signal. Moving to slightly longer response times, one encounters most veridical hits. For even slower responses, one eventually reaches the point where the misses and correct rejections are more common—as is known from the perception work of Treisman and Gelade (1980), a 'signal absent' response generally takes longer to decide upon than a 'signal present' one, especially when much distraction is present. The longest response times are more likely to be correct rejections than misses, on the assumption that for a typical correct rejection the detector will have more carefully considered more of the available information and used up more time, whilst an incorrect signal-absent judgement (miss) may occur due to an overly hurried judgement, perhaps stemming from the inverse of the 'gambler's fallacy' (it's got to *not* happen sometime). A hypothetical probability distribution of the four conventional SDT outcomes, as a function of response time is illustrated later in figure 5.

### 5. An air-traffic control example

To illustrate these conceptions further, the example of air traffic control given by Masalonis and Parasuraman (2000) is expanded upon. The seriousness of a conflict between two aircraft can be expressed in terms of the number of miles apart that the aircraft will be when they are at minimum separation. The distance apart is not the only thing that must be considered in the decision as to how worthy of attention a conflict is. The temporal factor is very important as well: if the aircraft are closing at a nearly head-on angle, then less time is available to deal with the situation. Conflict probes such as CPTP (Conflict Prediction and Trajectory Planning; Paielli and Erzberger 1997) take this into account by listing predicted conflicts in the chronological order in which they are predicted to occur. However, the ultimate area of concern is how close the aircraft are going to approach. There is some objectively true value of this minimum separation, but, due to factors such as winds and pilot actions, this value can not be known with 100% certainty until the actual time when the minimum separation occurs.†

Therefore, one can make a graph illustrating the distribution of the possible final separation values and the probability of each being correct. The distribution will change with time until, at the time of minimum separation, it approaches or becomes a vertical line at the  $x$ -value representing the minimum separation.‡ For an example,

---

† The issue of the certainty of measurement is ignored for now. According to Heisenberg, fundamental position cannot be measured with complete accuracy, at least not if one is also to know momentum. Accuracy of position measurement goes beyond the worlds of particle and quantum physics. In ATC, for example, the radar used to determine aircraft position possesses margins of error that are part of the reason behind the seemingly conservative separation standards (e.g. 5 nautical miles of lateral separation in cruise flight in the US) used in ATC.

‡ 'Approaches or becomes', depending on whether perfect measurement is assumed; see above footnote.

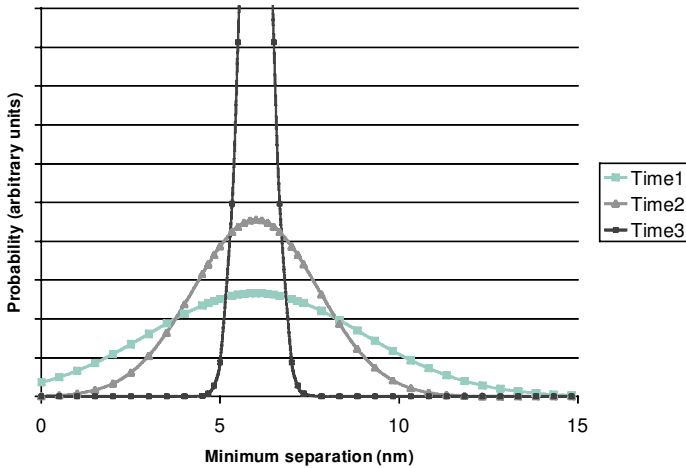


Figure 3. Distribution of predictions by a hypothetical conflict detector as to separation values between two aircraft that will experience a minimum separation of 6.0 nm.

take figure 3, which presents the likelihood that a hypothetical conflict detector will report a final minimum separation of various distances, given that the actual separation will be 6 nm. At early times (times 1 and 2, for example), the distribution of likely values for minimum separation is wide; certainty increasing as the time of the event approaches.

### 5.1. Sensitivity and bias

In analysing the performance of a detection system, Fuzzy SDT can be used to simply take a ‘snapshot’ of how accurate the detection was at the point the decision was made, by comparing the fuzzy decision at that given time to the actual value for separation. Over many trials with an equal value of separation and an equal amount of look-ahead time, a distribution similar to one of the above time distributions might be seen. By determining, in a fuzzy fashion, how much of a Hit, Miss, *FA*, and *CR* the response generated for each trial, measures can be derived of how sensitive and how liberal the detector was at the given time. There is no reason, by the way, that one would have to do this for only one value of separation; in fact the beauty of Fuzzy SDT is that one could have the detector make predictions in situations that would resolve to a number of different values of separation, assess its prediction each time, and derive versions of the standard measures of sensitivity and bias. If Fuzzy SDT parameters were calculated at a number of look-ahead times, one would expect the sensitivity—of an ATC conflict tool or indeed of *any* detection system predicting values of an analogue parameter—to increase as the time of the event approached, as the conflict detector became less of a 4-dimensional telescope, as it were, and more of a yardstick placed flush (in the temporal dimension) against a 3-dimensional situation that is now being more ‘measured’ than ‘predicted’ (see Brunswik 1955). The time-based effects on bias would be harder to predict and would depend on a number of factors. The way that the temporal dimension turns a ‘snapshot’ into a movie has been discussed. One can take a movie with traditional SDT, but a very dull and uninformative movie would result as the screen would suddenly go from white to black or vice versa (if in fact it changes at all). This is

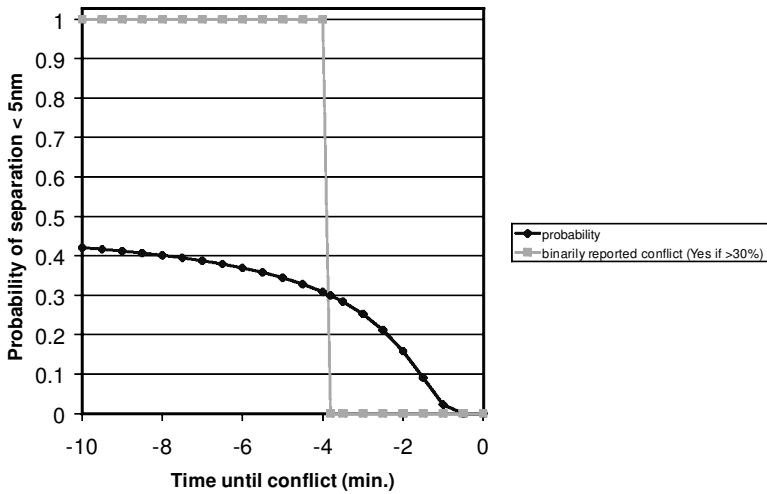


Figure 4. Distribution of predictions by a hypothetical conflict detector as to whether or not a conflict will occur between two aircraft that will experience a minimum separation of 6.0 nm, which is just over the legal separation minimum.

illustrated graphically in figure 4, which represents a hypothetical ATC conflict detection case, where the actual final separation between two aircraft will be 6 nautical miles. The detector's uncertainty about the actual separation—as illustrated in figure 3—starts large and shrinks as time goes on. Therefore, at early times the probability that the separation will be less than the 5 nautical miles criterion is some non-zero value. If the true separation is 6 nautical miles, the probability of a separation below 5 nautical miles decreases smoothly as the time of closest approach draws near. Assuming that the alerting function is liberal (prone to report conflicts) and binary, and reports a potential conflict if the probability that the aircraft will approach within 5 nautical miles is greater than 30%, then in this case the user sees a 'yes' judgment that a conflict will occur until time  $T-4$  when the detection system changes to report no conflict.

### 5.2. Alternative signal and response definitions

In Parasuraman *et al.* (2000), it is suggested that the value of 's', the extent to which an event contained a signal, and the value of 'r', the extent to which a response was made, must map into the range from zero to one. This enables a more faithful mapping of Fuzzy SDT analysis onto the mathematics of traditional SDT. However, other means can be proposed for analysing these situations. In any application where the variable of interest is definable along a continuum, the progression from the objective 'truth' to the response and the result thereof can be thought of in terms of a series of continua. First, there is the 'truth' continuum, and one will continue with the example of air traffic control. The seriousness of a conflict between two aircraft can be expressed in terms of the number of miles apart that the planes will be when they are at minimum separation. The detection system must make an accurate judgement of where along this continuum the truth lies. It will then report the truth as it sees it. The report might involve a judgement of where along the continuum the

truth actually lies. Alternatively, it may be a binary judgement about whether there will be a loss of separation, or what action the controller should take.†

The process can become a bit complex in the case of a multiple-detector system, because the output of one system component—the response—serves as the input—the signal—for the next component. In air traffic control, the objective truth, the state of the world, is that plane x and plane y will come within exactly 6 miles of each other in 8 minutes. The value of 6.0 miles is one exact point on the continuum and, after the fact, can be known as exactly the right answer (ignoring measurement error). The conflict detection automation analyses the flight paths of the two aircraft and might determine a number of things. The parameters it would estimate might include the most likely separation distance:

- (1) The most likely separation distance.
- (2) A distribution of possible separation distances and the probability of each.
- (3) A probability that the legal limits of 5.0 nm, or some other cut-off point will be violated.
- (4) A confidence judgement of #1 and #2.

In the case of #1 and #2, the reported distance is a point along a continuum and there may be little value in forcing it into a [0, 1] range just so that traditional analyses can be done. In case #3, there is more precedent for mapping a confidence

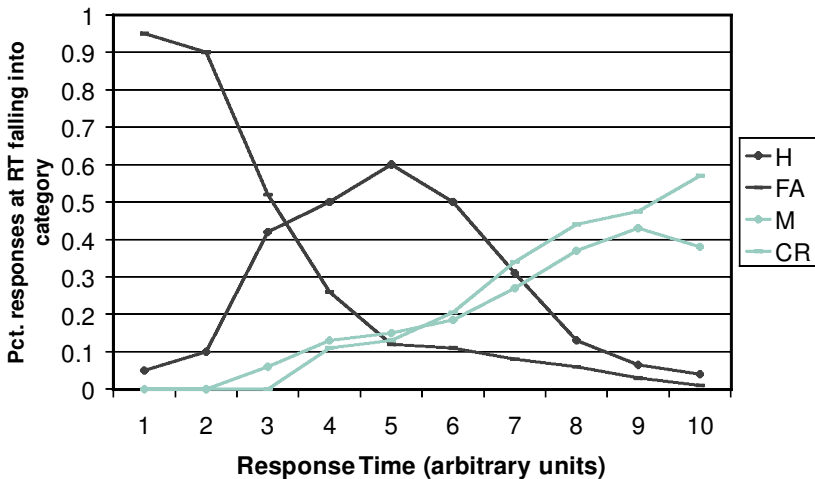


Figure 5. Proportion of responses at each level of response time that is associated with each of the four traditional SDT categories.

† At some point in many decision processes, as discussed in Parasuraman *et al.* (2000), a fuzzy input must be mapped to a discrete response, known in fuzzy logic control as defuzzification. The present discussion makes the point that discretization of a signal-response situation can occur at any point in the process. This can be done at the very start of the process, i.e. the definition of the signal can be collapsed to a binary state, or can be done in the next step, i.e. the judgement reported to the user (in this case the controller) is binary (in this case ‘conflict/no conflict’), or later in the process. It is the authors’ contention that if binary clipping must be done, the later it is done the better (i.e. at the stage when the final response is made).

judgement to a  $[0, 1]$  space, but it is arguable whether this is a true interval scale: does a 'confidence' of 100% mean twice as much confidence as 50%? What does it even mean to say 'twice as much confidence?' A probability (#4), however, maps easily into the range between zero and one.

An advanced automated detector could compute all these parameters and more for every pair of aircraft in the area without much trouble (Paielli and Erzberger 1997). The question then becomes how to report this information to the next component in the system—here, the controller. There are many possibilities, but what one is concerned with for the present discussion is the degree to which the automation's report of the 'truth' is fuzzy. It may be binary ('there will be a conflict' is a binary report, and even 'there is a greater than 80% chance that there will be a conflict' is a binary report). In such a case, the signal that the controller detects—the report of the automation—is binary, but this need not mean that the process must cease to be fuzzy at this point. Other information is available besides the report of the automation, such as the radar display of aircraft locations and directions, and the controller's memory of past situations and how the current situation compares with them. The controller may, therefore, use the binary fact that a conflict has been reported, and integrate this with the other data (s)he has available, in order to come up with a 're-fuzzified' judgement of the probability that a conflict will occur, or a value representing the eventual separation distance (potentially in zero-to-one space but quite possibly in actual units). This, in turn, will affect how quickly a clearance (instructions to change flight parameters such as course or altitude) is given, if at all, and perhaps with what degree of urgency it is communicated to the pilot, who, in turn, will make or fail to make a response based upon that signal. This process can be a closed loop, since pilot responses to clearances from the ground may, in turn, by their expeditiousness and the perceived success of the manoeuvre, act as signals affecting controllers' decisions regarding another kind of response, i.e. whether and to what extent they should monitor the cleared aircraft for conformance to the clearance. Pilot responses to clearances not only bear on the present situation, but also can create new conflicts—new signals to detect. Such a cascade effect is especially possible in the crowded airspaces which the ever-increasing demand for air transport is fostering.

## 6. Additional analytic frameworks

### 6.1. Redefining signal and response

It was questioned in the ATC example how much value there was in forcing a parameter that varies along a continuum in a  $[0, 1]$  range just to perform conventional SDT analysis. To handle these kinds of cases, where a  $[0, 1]$  mapping might be excessively arbitrary, it may be desirable to use different methods of analysis. One possibility is to use the mean and/or variance of the error score between the predicted and actual values on the parameter being detected. A mean negative error score (e.g. consistently reporting 5 nm for 6-nm separations) would show a liberally-biased system,<sup>†</sup> whilst a positive error score would show conservatism. The absolute value of the mean error score (or the variance of the response if the 'truth' or goal were always the same, as in RMSE as a measure of tracking performance) would be

---

<sup>†</sup> That is, liberal in terms of reporting conflicts that were more serious than they actually were.

a measure of sensitivity, and its variance would be an additional measure of predictability of the system; i.e. would its sensitivity and bias be consistent. A multiple regression of predicted values on actual values would accomplish more or less the same thing and, whilst not directly an SDT analysis and hence not directly comparable to traditional SDT in the way that the Fuzzy SDT methods originally outlined are, would provide a richer set of results than the output of a traditional SDT analysis.

### 6.2. *Applying fuzzy SDT to the cycle of automation design and use*

One could criticize that, in real world applications, the fuzzification inherent in a detection systems' confidence in its judgement is irrelevant, because, in the end, it either does or does not act as though a signal is present. However, at least in the case of a human making a decision, self-confidence as to whether one is executing the correct response can have very important consequences. For example, a decision of which a person is more sure will probably be executed more quickly—for better or worse. In addition, if communication with another agent is involved, the person's level of confidence may impact the result. In the case of ATC, a controller who decides that a command must be given to a pilot in order to avert a potential conflict may use specified phraseology (e.g. 'immediate,' or 'expedite'), or may issue the command in a more authoritative tone of voice if (s)he is more certain that a conflict or collision will result without intervention. Since the response of the controller to the potential conflict becomes, in turn, a signal to the pilot, the perceived confidence or urgency of the controller's command may affect the amount of time the pilot takes to execute the clearance, or may even affect the pilot's decision whether or not to question the clearance.†

The foregoing discussion of the cyclical, closed-loop nature of many signal detection tasks has omitted one step in the process. This has been purposeful in order that one might close the 'loop' at this point in the discussion. The heretofore-omitted step refers to the point at the beginning of the task, when a decision is made as to how a

---

† The same dynamics are at play on the battlefield and in the office. A commander or supervisor who gives an order has been taught to always project confidence. In effect, this results in a 'squashing' function where, no matter how sure the superior is that something needs to be done, (s)he may always communicate it with the same level of authority. This may argue that a better commander is one who uses a wider range of confidence when giving orders (perhaps always staying in the upper range, but varying across it). More to the point, it also has implications for automation design. Although the analogy between a human user of automation and a human supervisor of another human is often made, when one begins to speak of decision making or decision aiding automation, it may be more accurate to conceive of the automation as the supervisor. It has more information about the parameters that go into a necessary decision, but might integrate the parameters and derive a final answer that it provides to the human. Whether it somehow represents its confidence in its own decision is a question to consider. The past behaviour of the supervisor, the controller, the automation—has (s)he/it been right in the past?—will likely be combined with perceived or reported confidence in order to decide whether or how expeditiously to execute the suggestion or command. These considerations illustrate how the degree of trust vested by the user in himself or herself and in the automation (see Lee and Moray 1992, Masalonis 2000) combines with other factors, such as the reported confidence of the judgement, to produce a response decision. Of course, as discussed throughout this section, the trust, reported confidence, and final decision are all amenable to fuzzy representation.

signal should be defined. Here, one is referring to configurable decision-aiding automation, where the operator can set a threshold(s) that predetermines what states of the world will cause a response by the automation. Fuzzy and/or binary decisions can enter into the situation at this phase. For example, the ATC conflict probe known as CPTP allows controllers to specify cutoffs of lateral and vertical separation beyond which the automation will not report conflicts. This is an example of taking a variable defined along a continuum, viz. separation, and being permitted to set a cutoff at any reasonable point along the continuum—a cutoff that will cleave all states of the world into either ‘conflict’ or ‘no conflict’.† The ‘story’ of CPTP use, from beginning to end and back, goes as follows.

- The designer allows the user to specify a cutoff, a hard line beyond which all states of the world will be binarily assigned to a signal value of zero.
- The user (controller)—based on past experience on the job and perhaps with this automation— sets the cutoff along a continuum (say, 5–50 nm).
- The automation looks at states of the world that vary along a continuum (separation between each pair of aircraft).
- The automation makes a judgement about future conflicts along that same continuum and also makes judgements about each conflict along the continuum of probability (0–1).
- The controller receives this information. At this stage, it is assumed that the controller receives a 100% accurate representation of the automation’s judgements through the human–computer interface. That is, the interface and environment are such that the controller sees or hears the automation’s decision, and correctly interprets its meaning. There is no loss of information or fuzzification from the previous stage.
- The controller makes a binary but fuzzifiable decision. The decision is binary in that the controller either does or does not issue a clearance, and is fuzzy in that a greater or lesser amount of time will elapse before the clearance is given (due to the degree of certainty about the conflict and the degree to which this particular potential conflict is prioritized relative to other situations and tasks). It is fuzzy in that the clearance may be given with more or less urgency (due to the degree of certainty or perceived seriousness).
- The pilot’s response does or does not meet an end goal (prevention of an operational error), and can be said to meet or fail to meet that goal with varying degree.
- In turn, the result of the ‘trial’ may affect future decisions by the controller as to where to set the criterion in the future (or how much to trust the automation and how much to use it in the future; Lee and Moray 1992, Masalonis 2000).

A circumstance which tends more towards a binary situation is found in some radar detectors for automobiles, which allow the user to set a switch on either ‘city’ or ‘highway’, the former having a higher  $\beta$  in order to reduce the false alarms caused by the plethora of objects found in urban settings. In this example, the automation designer has set two cut-off levels between which the user can choose. Each of these

---

† It should be reiterated that the CPTP conflict probe, while necessarily using some user-specified cutoff for which situations to report as conflicts, does provide some fuzziness in its responses, by presenting the actual value for predicted separation rather than simply yes/no, and if controller desires, presenting the probability that a conflict will occur.

will divide all possible levels of the evidence variable into ‘object’ or ‘no object’, but some intermediate evidence levels will be classified as ‘radar’ by the ‘highway’ setting but as ‘no radar’ by the ‘city’ setting. It has been suggested (Lehto *et al.* 1998) that a useful feature on radar detectors would be a dial permitting users to vary the detector’s sensitivity along a continuum. The output would still be binary. Therefore, the progression of events would be as follows.

- The automation designer decides to make the criterion configurable along a continuum (high-to-low sensitivity).
- The user selects a point on that continuum.
- This point then serves as the cutoff point for the detector’s binary judgement (object present or absent). It should be noted as an aside here that the user’s car may be in the police’s radar beam to different extents, affecting the accuracy of the radar itself.
- The binary judgement of the radar detector tells the user about a binary ‘truth’ or state of the world (object present or absent). As in the ATC example, one assumes no loss of information or fuzzification from the previous stage: the driver receives a 100% accurate representation of the detector’s binary judgement regarding radar presence/absence (hears and correctly interprets the alarm).
- The output of the detector (along with other information available to the user) leads to a binary but fuzzifiable decision. The decision is binary in that the driver either will slow down or will not. The decision is fuzzy because the driver may slow down with different latency and different acceleration rates, decisions based in part on their confidence in the presence of radar and assessment of the degree of discrepancy between current velocity and desired velocity.
- The environmental result—that is, the success and timeliness of the action of braking leads to a non-binary (but basically categorical) level of success at achieving the end goal (the result is either ticket/no ticket, or warning).
- As in ATC, the results of a ‘trial’ lead in turn to future decisions about criterion setting, automation trust and automation use.

## 7. Wider implications

### 7.1. *Scientific phenomena as signals*

Having examined the foundations of Fuzzy SDT and its application, the discussion is now broadened to include wider implications of the concept that has been advanced. States of the world are not unambiguous; categorization is necessary if one is to accomplish one’s goals. The purpose of techniques, procedures and theories such as SDT is to provide ways of helping distinguish differences. This is undeniably helpful, and the authors are the last to suggest otherwise. However, in order to use such techniques, one must initially have a method of deciding into which group, signal or noise, to place each observation. This might seem facile, since instruments of measurement are often used for such purposes and, in today’s computer-driven world, one can easily become oblivious to the fact that such an initial decision must be made. Where the fact of this initial process is not recognized, subsequent findings and outcomes can induce a sense of confidence and certainty which is unjustified. The realm of research physics provides useful and important examples. In modern research in advanced physics, progress is crucially dependent upon sophisticated discrimination of differences at the edge of possible observation. In both celestial

and quantum mechanics, empirical observation occurs at levels of discrimination which would stagger the scientists of only a generation ago. Although the principle holds true for contemporary work, perhaps an example from early physical research best illustrates the point.

Whilst trying to polarize the newly discovered x-rays, the well-known French physicist Blondot claimed to have discovered a new form of radiation, named after his city of residence. Using prisms and lenses made of aluminum, Blondot claimed to be able to observe a spectrum of these n-rays by passing through it a fine thread coated with fluorescent material. Many individuals also claimed to see this phenomena, which was at or close to the then limits of observation. For a number of reasons, n-rays were eventually ‘debunked’ as unreal (Randi 1982, Youngson 1998) and they are no longer discussed in the research literature in physics. However, the combination of pronouncements of an admired individual and observations close to detection limits led to this unfortunate misidentification. It is all too easy to ‘blame’ the observer/discoverer here, but this is to be anachronistic and to fail to understand the milieu of research physics around the turn of the century. The central point of the story is that at or near discrimination limits human detectors, in this case trained scientists and observers, can ‘fool’ themselves into believing they have detected phenomena which actually are not real. The same is true of machine detectors programmed by humans.

### *7.2. Direction and magnitude of temporal s-r relationships*

Within the framework of temporal analysis, it is interesting to note that either fuzzy or traditional SDT can be and is applied to situations with different temporal relationships between signal and response. This variety of situation can be classed into ‘what was, what is, and what will come’, and can be explained most clearly through examples. Consider signal detection applications in the criminal justice system. Here, the signal might be represented by something that has already occurred—a crime. It is after the fact that the justice system attempts to determine whether a given suspect is responsible, using information subject to decay (e.g. loss of physical evidence, memory decay in eyewitness accounts). In contrast is the aircraft conflict detection example discussed above, in which future events are to be considered. This task, whether performed by human, machine, or both in tandem, attempts to derive a response regarding the degree of presence of a signal which has yet to occur. This type of detection, the task of prediction, faces its own set of challenges. Between the time the prediction is made and the time the signal actually occurs—or more congruently with Temporal Fuzzy SDT principles, the time when it is known with maximum certainty the extent to which the occurrence is a signal—any number of factors beyond the control or even the knowledge of the detection system might intervene to change the environment about which prediction was made.

As captured in figure 3, the further in time the signal and response are apart, the more uncertainty there will be regarding the degree of signal presence. This is true whether a response action anticipates or follows the signal. Therefore, the best predictions on average should be made when the signal and the response are essentially concurrent. They can not be exactly concurrent because of fundamental lags. For example, in the radar detector example, a ‘yes’ conveys information about what has already happened. Some systems have detection lags in the order of microseconds and so can appear to the observer to be virtually instantaneous. Other systems have much longer lags and affect behaviour accordingly. Assuming that

'what is' equals 'what was' is not always a good idea. For example, the last battle of the US Civil War was fought in the very southern-most part of Texas, many weeks after the official cessation of hostilities, illustrating that communication lags are nearly always crucial concerns. Most relevant information today travels much more quickly, and response times often need to be comparatively fast. Figure 3 has the same implications whether the scale of the base axis is nanoseconds or centuries, and whether the signal and response are really occurring at different times or whether the time discrepancy is caused by an information lag. The analysis methods proposed here can capture the dynamic nature of both perception and truth, regardless of time scale.

As has been noted, SDT is an exceptionally useful evaluative procedure, but it is certainly not the only one used by investigators in evaluating detection and decision-making performance. There is a parallel between SDT and another form of detection procedure which is used even more frequently; analysis of variance. For the purpose of exposition here, one refers only to a one-way analysis of variance (ANOVA) or more properly Student's *t*-test. Telling one thing from another is not a problem unique to psychological investigation. A student (William Gossett) was at one time employed by the Guinness Brewery to, amongst other duties, help ensure the quality control of their beer. His problem was to ensure that one sample of beer did not vary significantly from its companions, a problem much appreciated by many beer drinkers. In the course of his duties he proposed a statistical procedure to compare the difference between the means and the variability within each sample to decide whether a criterion difference between the two had been violated. Today, many students learn the nature of Gossett's procedure along their way to more sophisticated methodologies. SDT is directly related to the classical statistical procedure of one-way ANOVA, since each is concerned with the degree of similarity and difference between a bi-modal distribution of signal and noise or, more generally, two discrete samples. Simply put, both SDT and ANOVA are tied to at least one binary answer. In the former case, the input to the analysis consists of two black-and-white statements; signal present vs. absent, and response, yes vs. no. In the latter, we analyse a set of data points, each of which are definitively assigned to one and only one of two or more groups, then derive a yes-no answer regarding whether the groups are or are not drawn from the same sample and/or population. Whilst these two different techniques perform different transforms upon this data, in principle their actions are highly analogous. For example, if a gun is let off in a quiet room on several occasions, a human observer of normal hearing should provide a 100% hit rate and correct rejection rate, together with a 0% miss and false alarm rate. Whilst this person would be viewed as a 'perfect' observer, the self-same response pattern could be had by taking an artificial recording, such as one channelled through a microphone, and subjecting the derived scores to a one way ANOVA. In further analogy, when distinctions are difficult and errors are made, the categories of false alarm and miss are directly linked to statistical type I and type II errors, respectively.

Given this, can one devise a way in which the fuzzy nature of detection and selection in the first place can generate a method to help improve discrimination of events at a near threshold level? The authors believe that one can. In this respect, a fuzzy analysis of variance (FANOVA) is proposed, in which the degree of membership of a sample member is not a discrete situation but one which is characterized by a greater or lesser truth value. SDT provides a single 'snap-shot' of the situation at

hand, thus ANOVA and its various relations and derivatives are also largely bound to comparing situations at a moment in time (see also Ishibushi and Tanaka 1992, Ishibushi *et al.* 1993, Klir and Yuan 1995). However, just as the definition of a signal can change, so the membership value of a sample can vary over time, either as the sample itself changes or the criterion adopted for sample designation changes. ANOVA typically derives a probability value in which an arbitrary level of chance (i.e. 5%) is adopted as the criterion to accept or reject the notion of difference. FANOVA can be much more dynamic than this in its outcome.† Like a sequence of repeated analyses, which would be time-consuming and tedious in themselves, FANOVA could show how phenomena evolve over time.

The very word statistics comes from a semantic root meaning static or stationary. Thus, by entomological origin, it implies a single moment in time. Yet, behaviour is fundamentally characterized by variation and change, and, thus, time, in and of itself, is a crucial component of any full understanding. As one grows more sophisticated in various conceptual approaches to understanding, one must develop methods and procedures which match that level of theorizing. For the last century, statistics provided a crucial step forward along the methodological path. Now, in the new century, one must specifically and intentionally include time in one's efforts, and it is suggested that FANOVA can provide a first and small step along the path to dynamic statistics or 'dynistics.'

## 8. Summary and conclusions

One can arbitrarily reduce the world to black and white, but, in reality, it is not constructed that way. Although there is a great advantage in categorizing and sorting and, indeed, many aspects of cognition and development would not be possible without this capacity, the ability to appreciate the blend of items, objects, concepts and ideas also has great value. In this work, one has taken one of the pillars of human behavioural research and, indeed, detection assessment in general and elaborated it from the collapsed (black and white) case to show that the addition of some degree of uncertainty actually renders value. This technique can be applied to many different realms. Indeed, it is crucial to all of science, since the symbiotic acts of conception and observation constitute the central thrust of the scientific enterprise.

In an applied realm, the conception of fuzzy false alarms certainly has value. Traditionally, false alarms are viewed in their sterile form as an 'incorrect' response and in the practical realm as a nuisance to be suppressed when possible and muted where necessary. However, when a fuzzy false alarm is considered as a 'near miss', as occurs in aviation or ground transportation, one can see that this type of event represents a very informative occurrence, as can be illustrated by a final example from the aviation domain. The Aviation Safety Reporting System (ASRS) is predicated upon this assumption. ASRS is a compendium of reports of potential safety problems by pilots, controllers, and even airline passengers. These reports often stem from an incident where no damage or other impacts occurred but something nearly occurred. If considered in conventional SDT terms, these ASRS examples would be false alarms. In reality, the reports are responses to a 'close call' and, thus, are more

---

† One is aware of the linkages with correlation and with regression techniques. However, FANOVA can provide more information and additional advantages, although one has not attempted to articulate all such advantages here.

rightly viewed as 'near hits'. Many researchers have used this near-hit information as representative of trends indicating when actual collisions may occur. This form of practical application of a formal outcome of fuzzy signal detection indicates conceptually the value of considering membership functions beyond the traditional zero and one differentiation.

In case it has not been said clearly enough, one will now repeat, the authors are strong admirers of what has been called classic or traditional signal detection theory. It is their position that SDT represents one of the strongest quantitative techniques in existence for the analysis of human behaviour. Further, SDT can also be applied to machine detection and, by extension, to systems which use both human and machine in tandem in the detection process. Also, SDT has been used in many other realms as a useful and reliable technique. Thus, this work is not a critique of SDT, but rather stands in the hope of providing an extension to what is already a powerful ally. Through the joining with fuzzy set theory, one hopes to broaden even further the range of application and to use the advantages that fuzzy sets bring to provide more insight into behavioural response. In particular, it is believed that this marriage permits much further exploration and explanation of the temporal dimension of response in typical SDT situations and may, thus, further help in the endeavour to illuminate and understand difficult problems of behaviour in general.

#### Acknowledgements

We would very much like to thank Waldemar Karwowski and unknown reviewers for their helpful comments on an earlier version of the present work. This work was supported in part by Grant N00014-99-1-1034 from the US Navy to the first author.

Parts of this paper were produced under contract DTFA01-01-C-0001 to the MITRE Corporation, and are subject to Federal Aviation Administration Acquisition Management System clause 3.5-13, Rights in Data—General, Alt. III and Alt. IV (October, 1996). The views expressed reflect those of the authors and not necessarily those of the US Navy, the MITRE Corporation, the FAA or the US Department of Transportation. Neither the FAA nor the DoT makes any warranty of guarantee, expressed or implied, concerning the content or accuracy of these views.

#### References

- BRUNSWIK, E. 1955, Representative design and probabilistic theory in a functional psychology, *Psychological Review*, **62**, 193–217.
- CHIGNELL, M. H. and HANCOCK, P. A. 1986, A comparison of mental workload and available capacity in complex person-machine systems: a fuzzy-set theoretic approach, in W. Karwowski and A. Mital (eds), *Applications of Fuzzy-Set Theory in Human Factors* (Amsterdam: Elsevier Publishers), 271–288.
- COBB, L. 1981, Parameter estimation for the cusp catastrophe model, *Behavioral Science*, **26**, 74–77.
- DAVIES, D. R. and PARASURAMAN, R. 1982, *The psychology of vigilance* (London: Academic Press).
- FLACH, J., HANCOCK, P. A., CAIRD, J. K. and VICENTE, K. (eds) 1995, *Global Perspectives on The Ecology of Human-Machine Systems* (Mahwah, NJ: Lawrence Erlbaum).
- GARDNER, M. 1995, Fuzzy logic, *Skeptical Inquirer*, **19**, 9–11.
- GIBSON, J. J. 1975, Events are perceivable but time is not, in J. T. Fraser and N. Lawrence (eds), *The study of time II* (Berlin: Springer-Verlag), 295–301.

- GREEN, D. M. and SWETS, J. A. 1966, *Signal detection theory and psychophysics* (New York: Wiley).
- GUASTELLO, S. J. 1984, Catastrophe modeling of two opponent process models: drug addiction and work performance, *Behavioral Science*, **29**, 258–262.
- GUASTELLO, S. J. and MCGEE, D. W. 1987, Mathematical modeling of fatigue in physically demanding jobs, *Journal of Mathematical Psychology*, **31**, 248–269.
- HANCOCK, P. A. 1984, Environmental stressors, in J. S. Warm (ed.), *Sustained Attention in Human Performance* (New York: Wiley), 103–142.
- HANCOCK, P. A. 1993, Body temperature influences on duration estimation, *Journal of General Psychology*, **120**, 197–216.
- HANCOCK, P. A. 1997, *Essays on the future of human-machine systems* (Eden Prairie, MN: Banta).
- HANCOCK, P. A. and PIERCE, J. O. 1989, Integrating signal detection theory and catastrophe theory as an approach to the quantification of human error, *Paper presented at the Annual Meeting of the American Industrial Hygiene Association*, St. Louis, MO, May.
- HANCOCK, P. A. and WARM, J. S. 1989, A dynamic model of stress and sustained attention, *Human Factors*, **31**, 519–537.
- HUME, D. 1739, *A Treatise of Human Nature* (London: Noon).
- ISHIBUSHI, H. and TANAKA, H. 1992, Fuzzy regression analysis using neural networks, *Fuzzy Sets and Systems*, **50**, 257–265.
- ISHIBUSHI, H., TANAKA, H. and OKADA, H. 1993, An architecture of neural networks with interval weights and its application to fuzzy regression analysis, *Fuzzy Sets and Systems*, **57**, 27–39.
- JENSEN, D. D. 1989, Pathologies of science, precognition, and modern psychophysics, *Skeptical Inquirer*, **13**, 147–160.
- KARWOWSKI, W. 1992, The human world of fuzziness, human entropy and general fuzzy systems theory, *Journal of the Japan Society for Fuzzy Theory and Systems*, **4**, 825–841.
- KARWOWSKI, W. and MITAL, A. (eds) 1986, *Applications of fuzzy set theory in human factors* (New York: Elsevier).
- KLIR, G. J. and YUAN, B. 1995, *Fuzzy sets and fuzzy logic: Theory and applications* (NJ: Prentice-Hall).
- KOSKO, B. 1993, *Fuzzy thinking: The new science of fuzzy logic* (New York: Hyperion).
- KOSKO, B. 1997, *Fuzzy engineering* (New York: Prentice Hall).
- LEE, J. D. and MORAY, N. 1992, Trust, control strategies, and allocation of function in human-machine systems, *Ergonomics*, **35**, 1243–1270.
- LEHTO, M. R., PAPASTAVROU, J. D. and GIFFEN, W. J. 1998, An empirical study of adaptive warnings: human- versus computer-adjusted warning thresholds, *International Journal of Cognitive Ergonomics*, **2**, 19–33.
- MACMILLAN, N. A. and CREELMAN, C. D. 1991, *Detection theory: A user's guide* (New York: Cambridge University Press).
- MASALONIS, A. J. 2000, Effects of situation-specific reliability on trust and usage of automated decision aids, Unpublished Doctoral Dissertation, The Catholic University of America, Washington, D.C.
- MASALONIS, A. J. and PARASURAMAN, R. 2000, Application of fuzzy signal detection theory to analysis of human and machine performance in air traffic control, under review.
- PAIELLI, R. A. and ERZBERGER, H. 1997, Conflict probability estimation for free flight, *Journal of Guidance, Control, and Dynamics*, **20**, 588–596.
- PARASURAMAN, R. and DAVIES, D. R. 1976, Decision theory analysis of response latencies in vigilance, *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 578–590.
- PARASURAMAN, R., HANCOCK, P. A. and OLOFINBOBA, O. 1997, Alarm effectiveness in driver-centred collision-warning systems, *Ergonomics*, **39**, 390–399.
- PARASURAMAN, R., MASALONIS, A. J. and HANCOCK, P. A. 2000, Fuzzy signal detection: basic postulates and formulas for analyzing human and machine performance. *Human Factors*, **42**, in press.
- POWERS, W. T. 1973, *Behavior: The control of perception* (Chicago: Aldine).
- RANDI, J. 1982, *Flim flam* (Buffalo, NY: Prometheus Books).

- SORKIN, R. D. and WOODS, D. D. 1985, Systems with human monitors: a signal detection analysis, *Human-Computer Interaction*, **1**, 49–75.
- STEWART, I. N. and PEREGOY, P. L. 1983, Catastrophe theory modeling in psychology, *Psychological Bulletin*, **94**, 336–362.
- SUSSMANN, H. J. and ZAHLER, R. S. 1978, A critique of applied catastrophe theory in the behavioral sciences, *Behavioral Science*, **23**, 383–389.
- SWETS, J. A. 1977, Signal detection theory applied to vigilance, in R. R. Mackie (ed.), *Vigilance: Theory, operational performance and physiological correlates* (New York: Plenum).
- SWETS, J. A. 1996, *Signal detection theory and ROC analysis in psychology and diagnostics* (Mahwah, NJ: Erlbaum).
- SWETS, J. A. and PICKETT, R. M. 1982, *Evaluation of diagnostic systems: Methods from signal detection theory* (New York: Academic Press).
- TANNER, W. P. and SWETS, J. A. 1954, A decision-making theory of visual detection, *Psychological Review*, **61**, 401–409.
- THOM, R. 1975, *Structural stability and morphogenesis* (New York: Addison-Wesley).
- TREISMAN, A. M. and GELADE, G. 1980, A feature integration theory of attention. *Cognitive Psychology*, **12**, 97–136.
- WARM, J. S. 1984, *Sustained attention in human performance* (New York: Wiley and Sons).
- YOUNGSON, R. 1998, *Scientific blunders* (New York: Carroll & Graf).
- ZADEH, L. A. 1965, Fuzzy sets, *Information and Control*, **8**, 338–353.
- ZEEMAN, E. C. 1977, *Catastrophe Theory: Selected Papers, 1972–1977* (London: Addison-Wesley).

### About the authors

*Peter Hancock* is currently Provost Distinguished Research Professor in the Department of Psychology at the University of Central Florida. In his previous appointment, he founded and was the Director of the Human Factors Research Laboratory at the University of Minnesota. At Minnesota, he held appointments as Full Professor in the Departments of Computer Science, Electrical Engineering, Mechanical Engineering, Psychology and Kinesiology as well as at the Cognitive Science Center and the Center on Aging Research. He currently holds a courtesy appointment as a Research Scientist at the Massachusetts Institute of Technology (MIT). Dr Hancock is the author of over 200 refereed scientific articles and publications, as well as editing numerous books including *Human Performance and Ergonomics* in the Handbook of Perception and Cognition series, published by Academic Press in 1999 and the more recent *Stress, Workload, and Fatigue*, published in 2000 by Lawrence Erlbaum. He is the author of *Essays on the Future of Human-Machine Systems* released in 1997 and now in its third printing in CD format. He has been continuously funded by extramural sources for over two decades, including support from NASA, NIH, NIA, FAA, and FHWA as well as numerous State and Industrial agencies. In 1999, he was the *Arnold Small Lecturer* of the Human Factors and Ergonomics Society, and in 2000 he was awarded the *Sir Frederic Bartlett Medal* of the Ergonomics Society of Great Britain for lifetime achievement. He was the Keynote Speaker for International Ergonomics Association and the Human Factors and Ergonomics Society at the 2000 combined meeting in San Diego. His current experimental work concerns the evaluation of behavioural response to high-stress conditions. His theoretical works concern human relations with technology and the possible future of this symbiosis. He is a Fellow of and past President of the Human Factors and Ergonomics Society. He collects and studies antique maps and is a committed Ricardian.

*Anthony J. Masalonis* is a Senior Multi-Disciplinary Systems Development Engineer at The MITRE Corporation's Center for Advanced Aviation System Development, and supports the design and evaluation of concepts for the automation systems used by Traffic Management Coordinators (TMCs), Air Traffic Controllers, and other professionals in the aviation system. To these tasks he applies expertise in Human Factors/Ergonomics, statistical and database analysis, and training design/evaluation. In 2000 he earned a PhD with Distinction in Applied-Experimental Psychology from the Catholic University of America, where he studied trust in automation and how it is affected by variable reliability and user training, and also helped to develop the basic postulates of Fuzzy Signal Detection Theory. His theoretical

interests include merging the trust and SDT frameworks, along with other frameworks of automation theory, and considering how to apply the resulting chimeras to human–automation system development. He is a writer of poetry, music and fiction, has published creative works in several literary journals, and performs on the Washington D.C. poetry circuit.

*Raja Parasuraman* is Director of the Cognitive Science Laboratory and Professor of Psychology at The Catholic University of America in Washington D.C. He received a BSc (1st Class Honors) in Electrical Engineering from Imperial College, University of London, UK (1972), and an MSc in Applied Psychology (1973) and a PhD in Psychology from the University of Aston, Birmingham, UK (1976). He has carried out research on attention, automation, air-traffic control, aging and Alzheimer's disease, event-related brain potentials, functional brain imaging, signal detection, vigilance and workload. His research in these areas has been supported by several federal agencies, including the National Aeronautics and Space Administration (NASA), the National Institutes of Health (NIH), the US Navy, as well as by private foundations. His books include *The Psychology of Vigilance* (Academic Press, 1982), *Varieties of Attention* (Academic Press, 1984), *Event-related Brain Potentials* (Oxford University Press, 1990), *Automation and Human Performance* (Erlbaum, 1996) and *The Attentive Brain* (MIT Press, 1998). He was elected a Fellow of the American Association for the Advancement of Science (1994), the American Psychological Association (1991), the American Psychological Society (1991) and the Human Factors and Ergonomics Society (1994). Dr Parasuraman was Chair of the Human Performance and Cognition Study Section for the NASA Neurolab Mission in 1995, served as a member of the Human Development and Aging Study Section of NIH from 1992 to 1995, and was a member of the National Research Council's Panel on Human Factors in Air-Traffic Control Automation from 1994 to 1998. He is also the current Chair of the National Research Council Panel on Human Factors.