Modelling Decision–Making by Pilots

Progress Report

December 13, 1993

Nicholas J.M. Patrick

Human-Machine Systems Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

This work is sponsored by the NASA Ames
Research Center under contract number NAG2-729

Contents

1 Introduction

Our scientific goal is to understand the process of human decision-making. Specifically, we seek a model of human decision-making in piloting modern commercial aircraft which prescribes optimal behavior, and against which we can measure human sub-optimality. This model should help us understand such diverse aspects of piloting as strategic decision-making, and the implicit decisions involved in attention allocation.

Our engineering goal is to provide design specifications for (i) better computer-based decision-aids, and (ii) better training programs for the human pilot (or human decision-maker, DM).

2 On Models

2.1 Desirable Attributes of a Model

The following is a list of desirable properties of a model:

(i) Predictive Ability — A model should allow accurate prediction of the important behaviors of the modelled system[1]. This is the defining property of a model.

(ii) Generality — To be broadly applicable or *general*, a model must be based on the internal structure of the situation or object being modelled. General models should have the desirable property that new behaviors can be predicted with only *small* changes to the model, such as parameter changes.

(iii) Simplicity — Occam's razor, the principle that the simplest of competing theories should be preferred, suggests that simple models are more general. Because this is the opposite of complexity, it might best be measured using the notion of Kolmogorov Complexity[2]. Alternatively, the number of parameters which must be specified might be used as a measure of complexity.

(iv) Normativity — We desire a model which specifies "optimal" behaviour in the process being modelled.

---

1 This is possible with aircraft because they have relatively simple and well-understood dynamics. However, prediction of pilot behavior is difficult. However, the supervisory behavior of modern pilots is inherently nonlinear and complex, so that systems containing a human pilot will exhibit chaotic behavior. Complex models of this behavior will do the same, making prediction difficult or impossible. By this, we mean that although statistical predictions will be possible, for instance we will be able to estimate on what fraction of flights in given conditions a pilot will choose Y, we will never be able to predict the outcome of a particular decision at the end of a particular flight.

2 See Yufik, Y.M. & Sheridan, T.B.; A Technique to Assess the Cognitive Complexity of Man-Machine Interface. Philadelphia, PA; Institute of Medical Cybernetics. p. 23.

(v) Tractability — We desire a model which is susceptible to mathematical analysis[3].

A model which is general, simple, and tractable, might be said to be insightful: capable of conveying insight into the behavior of a system.

Neural-networks (NN's) fare better against some of these criteria than against others. Although excellent for prediction, NN's convey very little insight into the process being controlled (neither are they simple), or the nature of the decisions being implemented. Also, they do not have the first desirable property: generality — they cannot easily be tuned, they must be retrained.

Econometricians commonly use models which do not posses all of these properties: regression models of the economy are easy to construct and often have good predictive power, but they are usually not general.

2.2 Testing Models

2.2.1 Testing against an Objective Function

Humans are required in these strategic situations because the decision-making cannot be automated. Since we have no objective function with which to automate, we have no normative model against which to evaluate the <u>optimality</u> of the human subject's decisions. This has been called Roseborough's dilemma[4]. We don't know which choices people should be making because we can't compare the values of the choices without a utility function. Any such utility function must usually be derived from these same choices. This is a circular procedure.

2.2.2 Testing using Consistency

However, there are objective methods of testing the <u>predictive ability</u> of a model. They involve testing the following three types of consistency:

(i) Consistency within the training set — how well the model fits the data used to build it[5];

(ii) Consistency within the subject — how well the model fits extra data taken from the same subject (this can be extrapolative or interpolative);

(iii) Consistency across subjects — how well one subject's model predicts the behaviour of other subjects .

---

3 Rule-based models cannot do this well.

4 Roseborough, J.B. "Aiding Human Operators with State Estimates". Ph.D. Thesis, 1988, M.I.T., Cambridge, MA.

5 Taking into account the number of degrees of freedom available in the model — the fewer parameters, the more significant this consistency is.

We propose that the extent to which human or machine decision-makers are inconsistent is a measure of their sub-optimality. Several candidate measures of inconsistency might be used. Among them are (i) the number of errors made by an expert system, (ii) a measure of the lack of symmetry embodied in an expert system's decision rule[6].

A method which can be used to augment this approach is to take some <u>cardinal</u> subjective measures (e.g. the subjective degree–of–difference between the alternatives) as each decision is made by the subjects, and compare these with measures taken from the models. This technique should make assessment of the models easier than purely ordinal data (e.g. the alternative chosen) do.

3 Models of Decision-Making

3.1 Types of Decision

Philosophically, decisions may be divided into three types:

(i) Value-based [optimal] decisions — in which an objective [perhaps utility] function is maximized;

(ii) Rule-based decisions — in which a learned rule is applied without direct attention to the values of the possible outcomes;

(iii) Random decisions — these are not quite decisions, just random apparent choices.

The first two of these types, optimal and rule-based decisions, bracket a spectrum of decision models. It is worth noting that even though rule-based decisions employ no immediate notion of value (by definition), they are optimal only to the extent that they are learned from optimal decisions. Any theory of decision-making needs to take into account the concurrent use of each of these three decision-making strategies.

An interesting idea is that people learn <u>not</u> to make utility-based decisions, by learning simple rules which can be applied in their place. In other words, with experience they learn to substitute rule-based decisions for value-based decisions. This idea might be applied to the process of choosing between control strategies[7].

3.2 Applications for Decision Theory

We are pursuing a two-pronged approach to modelling human decision-making, with application to piloting. The first involves assessment of the values or features used in comparing the alternatives in a decision. The second involves the application of decision theory to a dynamic task.

---

6 The decision used in these experiments should be symmetric — that is to say insensitive to the order in which the alternatives are presented.

7 To demonstrate this we would need an experiment in which subjects can be shown to reject utility-maximization in favor of some other, specific decision rule.

3.2.1 Strategic Decision-Making

Of obvious interest in aviation is the application of decision theory to the modelling of strategic decision-making, in which pilots choose between goal alternatives. Here, value might be assigned to each of the alternatives, and a maximization scheme like Subjective Expected Utility (SEU) might be used to model the choice between them. The experiment described in Section 4 is designed to examine this process.

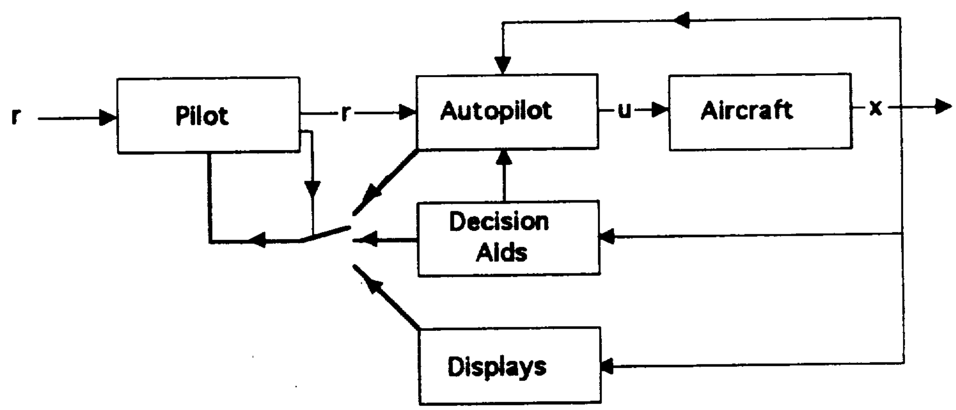
3.2.2 Dynamic Attention Allocation

We would also like a model of pilot's attention-switching behavior between control and monitoring tasks in a dynamic environment, as suggested in Figure 1. For this we might again invoke a utility-maximization model, like the explicit-cost model in Equation 1 below:

$$V^{*}_{net} = \sum_i p(x_i)\cdot \max_j\left\{V(u_j x_i)\right\} - \max_j\left\{\sum_i p(x_i)\cdot V(u_j x_i)\right\} - \sum_i c(x_i)p(x_i)\log_2\left\{\frac{1}{p(x_i)}\right\} \quad [1]$$

where x_i are the possible states of the world, $p(x_i)$ is the probability of occurrence of a particular state, u_j are the possible control actions, $V(u_j x_i)$ is the value associated with the application of control action u_j to situation x_i, and $c(x_i)$ is the explicit cost of one bit of information about state x_i.

Figure 1: The pilot controlling information and control automata, and sampling aircraft state directly. The competing resource sinks of autopilot and decision-aid attention allocation are shown by the heavy dashed lines.



However, this model requires a utility function on the *entire* space of possible outcomes x_i. This is hard to obtain [perhaps unattainable] in a flying environment, where the values and probabilities, particularly of the rarest and most extreme outcomes, are usually unknown, and never explicitly stated[8]. This is in contrast to prices in an eco-

8 So-called optimal control cannot be strictly optimal, since it relies on quite arbitrary cost-functions.

nomic market, which – because of trade between individuals – are usually available explicitly[9]. For example, in an economy, labor and capital are commonly exchanged between corporations, thus explicit prices for both must be made available. The in-flight decisions in which we are interested involve non-tradable attributes of the DM's situation. For instance, a pilot cannot trade his aircraft's altitude for the airspeed of another aircraft, and thus no market with associated explicit prices exists for these attributes. In this sense, flying involves situations, rather than goods[10].

One method of bypassing this problem would be to assign costs and rewards for all possible outcomes in an experiment. With this externally-imposed utility function, we would be able to examine the optimality of the subject's decisions, but not the process by which he values the alternatives. If we don't understand this process, we will not be much closer to understanding decision-making. We might try to elicit a utility function for decision-making from the subject, or we may use some other [feature-based] criteria in our model.

A decision model should shed some light on the process by which pilots allocate their limited attention resources between tasks in the cockpit. For instance, pilots must choose between controlling the aircraft through the autopilot and manipulating decision aids (such as the FMS or other navigation equipment) to provide information relevant to the control task. Figure 1 above shows these two competing tasks. A second experiment (described in Section 8) is designed to focus on this application of decision theory.

4 Strategic Decision–Making Experiment

In this experiment, experienced pilots were instructed to imagine themselves on a long cross-country flight in Instrument Meteorological Conditions (IMC) in a particular aircraft[11], when the destination airport became unusable for a non-weather-related reason. The conditions for the flight (e.g. no thunderstorms, unlimited visibility below the overcast, etc.) were specified as completely as possible.

Each subject was instructed to choose between about three dozen pairs of alternates based on the distance to, and the ceiling at, each alternate. For example, subjects might have chosen the second airport from the following pair: [*200 n.miles, 1000 feet*] vs. [*50 n.miles, 500 feet*]. The subjects were also asked to indicate the subjective "degree of difference" between the alternates on a numerical scale of 0 to 7. Each subject's utility was assessed[12] independently from the paired comparisons. Subjects were given a scale with anchor points ([worst ceiling, worst distance], and [best ceiling, best distance]) al-

---

9 Price is only a crude substitute for utility, because it reflects the averaging of utilities from many individuals, both large and small.

10 One way around this problem might be to create an artificial market for these attributes. In an experiment, several pilots might fly aircraft in different situations concurrently, and be allowed to trade these attributes over the radio.

11 The subjects were asked to imagine that they were in the aircraft they usually fly.

12 After Yntema, D.B. and Klem, L.; "Telling a Computer How to Evaluate Multidimensional Situations"; IEEE Transactions on Human Factors in Electronics, Vol. HFE-6, No. 1; June 1965.

ready marked, and instructed to mark on it the two intermediate points ([best ceiling, worst distance], and ([worst ceiling, best distance]) in such a way that the distance up the scale was proportional to the desirability [safety] of the alternate. They were then given subsidiary scales for distance and ceiling (again with anchor points) and instructed to mark the intermediate points on each scale in the same way.

Figure 2: The experimental situation, as presented to the subjects. The alternates are located so as to be equally inconvenient for travel to the ultimate destination.

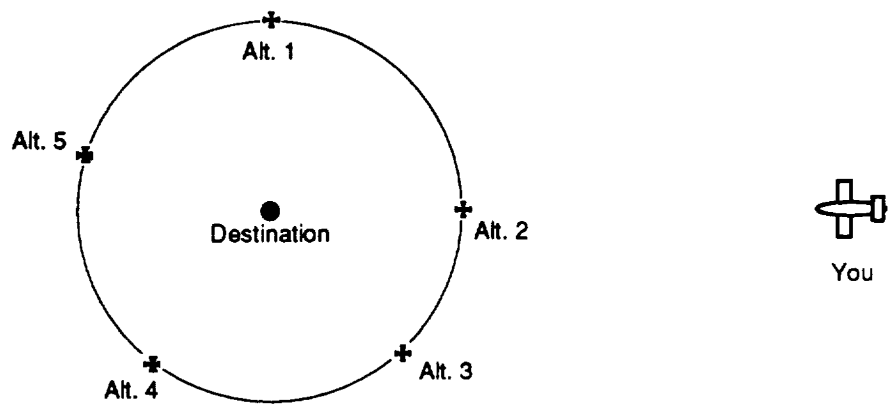
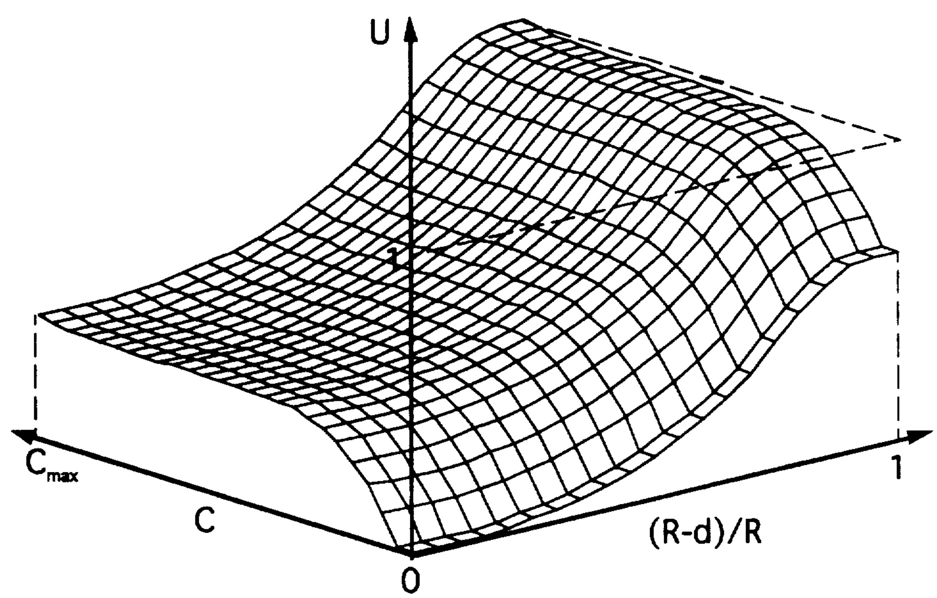


Figure 3: Utility as a function of the two independent variables of the experiment: normalized range-remaining and ceiling. [subject R]

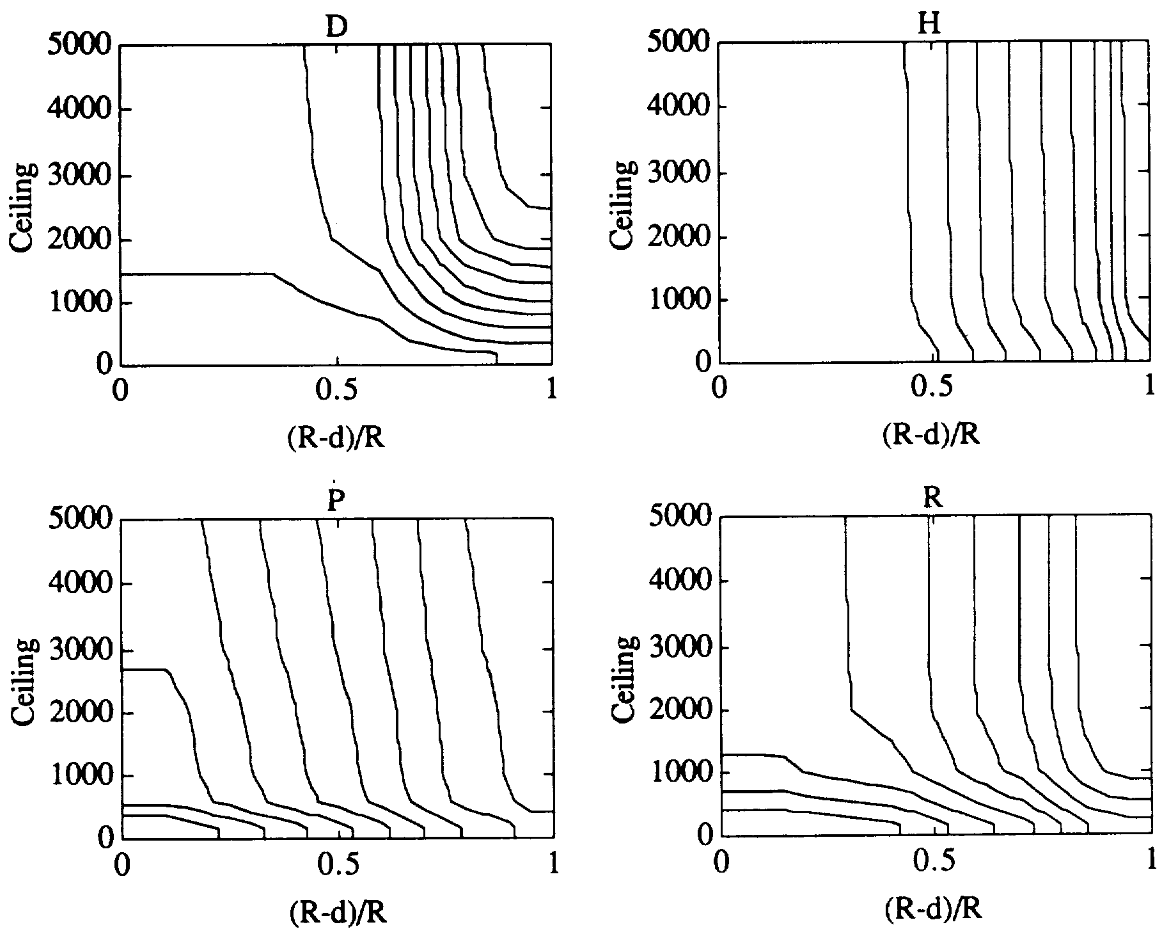


5 Decision Models

Several types of decision model were applied to the data extracted from the subjects in the experiment above. A utility model was applied to the decisions, and 3 types of discriminant–function model were applied to both 2– and 4–dimensional incarnations of the decisions.

5.1 Utility Model

Figure 4: Contour plots of utility for the 4 subjects as a function of the two independent variables of the experiment: normalized range-remaining and ceiling. In each plot, U=0 in the lower left corner, and U=1 in the upper right corner. Contour lines are at a spacing of ΔU=0.1.



A quasi-separable utility function, as defined by Keeney[13], was used to fit the corner points of the first ("corner-point") direct-assessment scale:

[13] Keeney, Ralph L.; "Evaluating Multidimensional Situations Using a Quasi-Separable Utility Function"; IEEE Transactions on Man-Machine Systems, Vol. MMS-9, No. 2; June 1968.

$$U(x,y) = a + bU(x,0) + cU(0,y) + dU(x,0)U(0,y) \quad [2]$$

with constants *a, b, c,* and *d* chosen so that *U* lies on the range [0,1]. The component utilities, *U(x,0)* and *U(0,y),* were measured directly from the points on the subsidiary scales, and linear interpolation was used between these points in evaluating the utilities of the alternates. The distance variable, *x,* was normalized with aircraft range.

This method of utility evaluation was used to determine utility as a function of both ceiling and distance variables. Figure 3 shows a typical utility function.

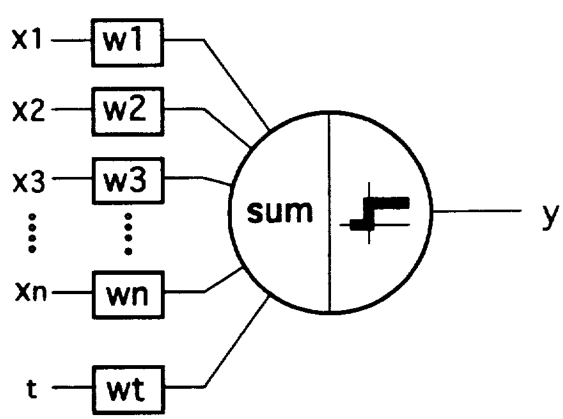
5.2 Discriminant–Function Models

It might be possible to discriminate between alternatives in a decision by using linear combinations of the alternatives' attributes, rather than subject-dependent non-linear functions of the attributes, as the utility method does. To examine this possibility, several linear discriminant–function models were tested on the data. Each attempts to fit a hyperplane (flat surface) between the two classes of decision points (those in which the first alternative was chosen, and those in which the second alternative was chosen).

5.2.1 Perceptron Algorithm

A perceptron is a very simple neural network. The perceptron algorithm is guaranteed to converge for small enough step sizes, if the data are linearly separable.

Figure 5 A Perceptron, or single-element neural network, which can be trained to find a discriminant hyperplane to fit the binary, forced-choice decisions.



The algorithm was written in LISP, and run with each of the subjects' decisions as training sets. To examine the possibility that the DM's might use fewer than the 4-dimensions available to them in each decision (r_1, c_1, r_2, c_2), the algorithm was rerun in 2 dimensions for each subject. The two dimensions used were:

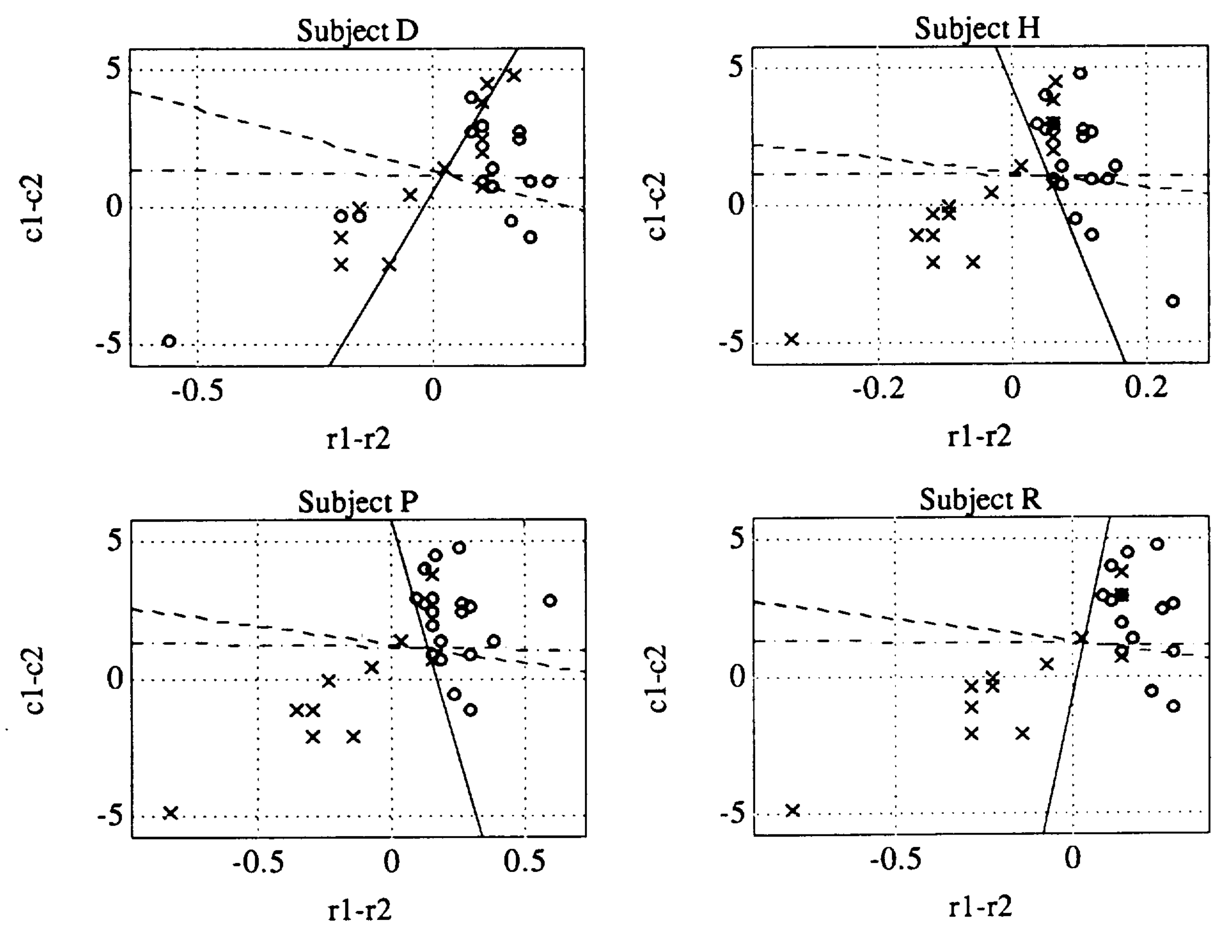
$$r' = r_1 - r_2, \quad c' = c_1 - c_2 \quad [3,4]$$

The discriminant functions obtained for these two-dimensional cases (see the explanation below) are shown as solid lines in Figure 6 below.

5.2.2 Closed-form Discriminant Algorithm

Observe in Figure 6 that the discriminant functions determined by the perceptron algorithm vary widely from subject to subject. This is because the algorithm is very sensitive to individual data points. In an attempt to find a more general model of the decision-making, a less sensitive closed-form algorithm was designed. The means and standard deviations (in each dimension) of both data groups were calculated, and a discriminant hyperplane was constructed so as to be normal to the line connecting the two means, and to pass through that line at some fraction along its length, determined using the standard deviations of the groups. These functions are shown in Figure 6 as dash-dotted lines. The large difference in the slope of the closed-form line and the slope of the perceptron line is explained below.

Figure 6: Plots of 2-dimensional decisions (reduced from 4-D by subtraction) for the four subjects. Difference in Ceilings is plotted against difference in normalized range. The perceptron discriminant functions are shown by the solid lines, the closed-form discriminant functions by the dash-dotted lines, and the improved closed-form discriminant functions by the dashed lines[14].



14 Subject D's responses are noticeably different from the others'. That subject was unsupervised for the duration of the experiment.

5.2.3 An Improved Closed-form Discriminant Algorithm

The above algorithm does not perform particularly well (see Figure 6 and Table 1), because it does not take into account the difference in spread of the data along the two axes. Since units of measure are quite arbitrary, any algorithm which relies on them for the scaling of its axes is likely to be biased. For example, if ceiling is measured in feet, rather than in thousands of feet, the importance of ceiling in the decision will be exaggerated by a factor of 1000. This problem would disappear if both axes were normalized to lie on the range [0,1]. Alternatively, the angle of the discriminant plane can be adjusted to reduce this bias. Figure 7 below shows why this improvement to the closed-form discriminant-function model works.

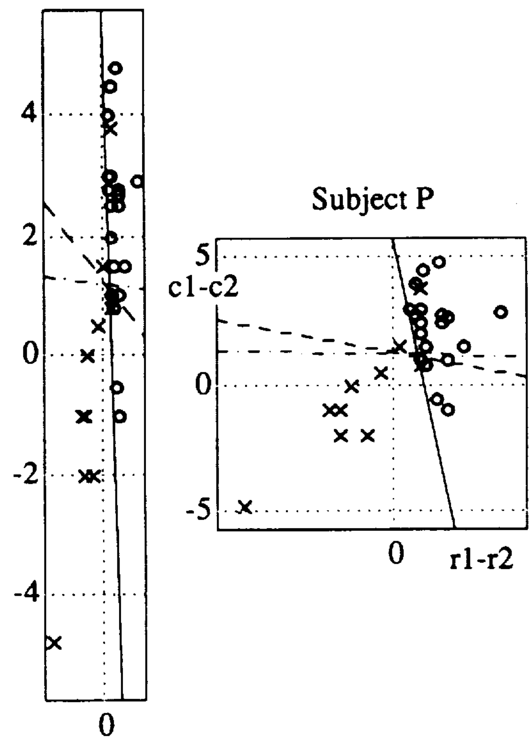


Figure 7: Two plots of the 2-dimensional decision. On the left, the decision as it appears if the axis are scaled in accordance with their ranges — values along the x-axis vary from -1 to +1. Since units of measure are arbitrary, this can exaggerate one dimension of the decision. On the right, the decision as it appears if the axes are scaled equally. This seems more realistic. The standard closed-form discriminant function is shown in both as a dash-dot line, the improved closed-form discriminant function as a dashed line, and the perceptron function as a solid line. Notice how the improved function separates the two cases more effectively than the original closed-form function.

This improved algorithm was run on the 2-D and 4-D data from each subject. The results are summarized in the tables below.

6 Comparison of Results

Following the argument - presented in Section 2.2 - that we have no objective measure of the quality of a decision, we will use consistency as a proxy for optimality.

6.1 Within-Subject Consistency

Table 1 below shows the percentage of decisions for which the expert system and the human subject did not agree. Note that we can not say "the percentage for which the expert system erred" any more than we can say "the percentage for which the human erred". All we can examine is consistency. First observe that for the 2-D and 4-D discriminant function models, the Perceptron algorithm produced the best results. This is not entirely surprising because it's only goal is the reduction of the *number* of classification errors[15], whereas the other models could be said to be reducing the *size* of the errors. Second, in the 2-D case, the improved closed-form algorithm ("Closed Form 2") did a better job than the original closed-form algorithm for every subject's data. Note, however, that for the 4-D case it did no better. This is somewhat surprising, and warrants further examination.

Table 1: Percentage errors made by several different expert systems mimicking the decision-making of the 4 subjects: utility data from Yntema-Klem style experiment; number (and percentage) of perceptron discriminant-function errors; number (and percentage) of closed-form discriminant-function errors; Experience (hours total time) and age (years). The subjects made [31,36,33,30] decisions.

Subject		D	H	P	R	Average
Utility Model		29 %	14 %	6 %	13 %	16 %
4-D	Perceptron	23 %	6 %	3 %	0 %	8 %
	Closed Form	42 %	33 %	12 %	23 %	29 %
	Closed Form 2	42 %	33 %	18 %	23 %	29 %
2-D	Perceptron	26 %	14 %	3 %	10 %	13 %
	Closed Form	48 %	36 %	21 %	27 %	33 %
	Closed Form 2	36 %	31 %	18 %	20 %	26 %
Total Time (hrs)		420	4200	280	480	1345
Age (yrs)		25	37	21	28	28

Third, note that the Utility model is not as good as the perceptron algorithms for any of the subjects[16], but is better than any of the other models for every subject. The results for the utility function are slightly biased because we used separate cardinal [ratio] utility data to obtain the utility function, whereas we only used the training-set data for the other models.

6.2 Between-Subject Consistency

Table 2: Within-subject and between-subject error rates of the seven models. Each model was applied to each data set without any parameter adjustment, except that in each case distance was non-dimensionalized. The ratio of the two rates (between/within) is shown.

Model		Within	Between	Ratio
Utility		15.6 %	23.6 %	1.5
4-D	Perceptron	7.8 %	22.0 %	2.8
	Closed Form	29.2 %	34.9 %	1.2
	Closed Form 2	29.2 %	34.3 %	1.2
2-D	Perceptron	13.2 %	21.0 %	1.6
	Closed Form	33.1 %	33.1 %	1.0
	Closed Form 2	26.1 %	27.6 %	1.1

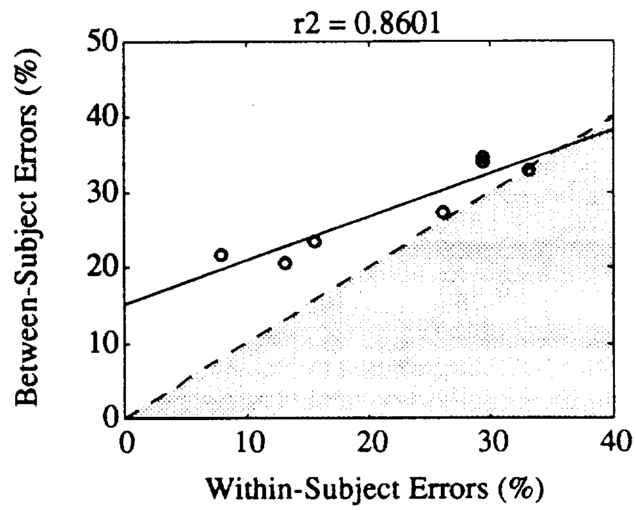
15 It might be argued that the value of the perceptron algorithm lies in its ability to provide a lower bound on the minimum number of classification errors.

16 If we were using a different statistic to measure the consistency this might not be the case. For instance, the number of errors weighted by the size (difference in utility) of each error might be more favorable to the Utility model. This is the measure Yntema & Klem used, but they were only comparing utility models with one another. We do not have a measure of "difference" which is consistent for both utility and discriminant function models, so we cannot easily compare the two with such a statistic.

In order to test the generality of the models, each person's models were used to mimic the decisions of the other subjects. The results appear in Tables A1-A7 in Appendix A, and are summarized in Table 2 below.

The ideal model would combine low within-subject error rates with low between-subject error rates. Figure 8 shows these error rates.

Figure 8: Average between-subject error plotted against average within-subject error for each of the 7 models. A regression line (solid) is shown, as is the line y=x (dashed). Models should not (except by chance) lie in the shaded region, where they would be better predictors of other subject's data. A completely general model would lie on or above the line y=x near the origin.



6.3 A Subjective Measure of Difference

The subjective "degree of difference" value recorded in the utility assessment experiment should be a measure of the ease of the decision between the two alternatives. This measure was plotted against difference in utility in the case of the utility model, and against distance from the decision plane for the other models. Surprisingly, there is little correlation between any of these measures. Efforts to correlate the subjective difference with the difference in distances alone produced similar results. This rather surprising result warrants further investigation (see Section 7.1 for a description of an improved experiment).

Table 3: Correlation coefficients for linear relationships between objective distance measures and subjective difference between the alternatives, using the following objective measures of distance: (i) for utility model, distance = difference in utility; and (ii) for discriminant function models, distance = distance from decision plane .

Subject		D	H	P	R	Average
Utility Model		-0.0336	0.0486	0.1217	0.0889	5.6 %
	Perceptron	0.0195	0.0202	0.2053	0.0154	6.5 %
4-D	Closed Form	0.0625	0.0859	0.2900	0.2265	16.6 %
	Closed Form 2	0.0264	0.0813	0.2625	0.2115	14.5 %
	Perceptron	0.0093	0.0074	0.0016	0.0157	0.9 %
2-D	Closed Form	0.0874	0.0212	0.2538	0.1219	12.1 %
	Closed Form 2	0.0630	0.0157	0.2166	0.0989	9.9 %

Note that the basic closed-form algorithm consistently produced the best results.

7 Future Work I

7.1 An Improved Utility-Estimation Experiment

It is possible that the experimental data obtained in the strategic decision experiment outlined herein were compromised by the lack of realism in the experimental setup. We feel that a graphical presentation of the alternatives might encourage more care and consideration in the responses. Accordingly, we are writing a portable version of the experiment to run on a Macintosh PC. The prototype of the display is shown below in Figure 9.

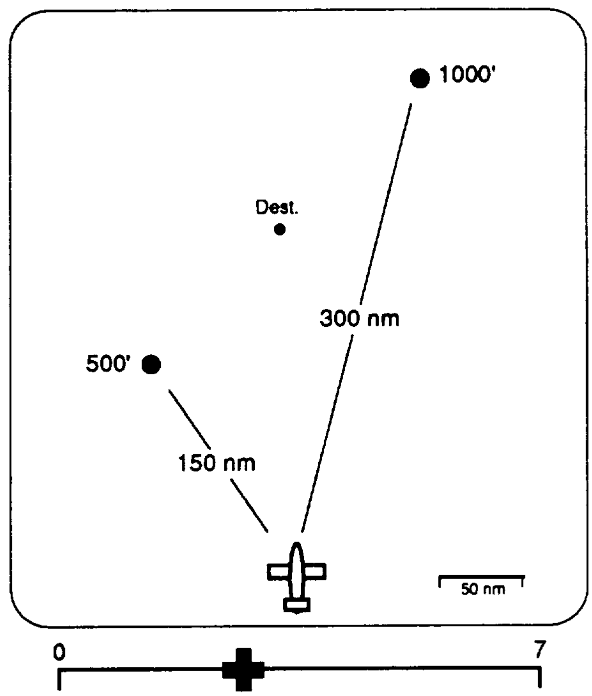


Figure 9: The display for the Macintosh-based Utility-assessment experiment. The subjects will be presented with a fresh window, like this one, for each decision. After considering the range and ceiling information presented for each alternate airport, they will click on the desired one, and will then use the mouse to place the pointer (+) at the appropriate location to indicate the subjective "degree of difference".

We expect that with this setup we will be able to obtain more accurate data, and may be able to examine the issue of the subjective "degree of difference" more carefully.

7.2 Improvements to the Analysis

The following might be done to further examine the data already obtained:

(i) Incorporate a measure of asymmetry into the discriminant function analyses — each of the algorithms should be insensitive to the order in which the alternates was presented (except for the direction of it's output). This was clearly not the case.

(ii) It would be interesting to perform Multi-Dimensional Scaling (MDS) on the decision data, but this must wait for better subjective degree of difference data, which we hope to acquire from the experiment outlined in Section 7.1.

(iii) The interpolation of the subsidiary utility scales was done linearly. Several alternative algorithms were tried on the data, but were unsatisfactory. For instance, when a cubic spline was fitted through the data, the resulting utility was no longer always monotonically increasing, and was not always positive for some of the subjects! A more interesting curve-fitting algorithm is required.

(iv) What does the utility surface look like in 4-D? It's meaning? Results?

(v) We would like a method of learning the utility function from the ordinal information contained in the paired comparisons, rather than from the cardinal information contained in the separate assessment used in our experiment.

(vi) We would like to be able to fill out the following table, using numerical measures of the degree to which each model possesses each property.

Table 4: Properties of various models of human decision-making.

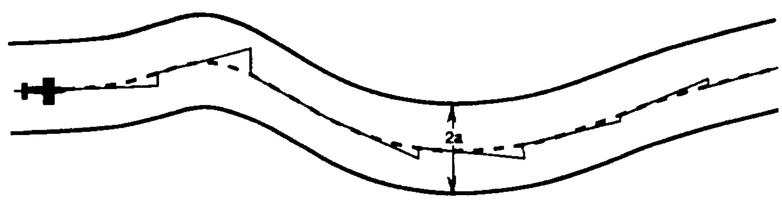
	Non-linear NN's	Discriminant Functions	Rules	Utility
General				?
Simple		√	√	√
Normative				√
Predictive	√	√		√
Tractable	√	√		√
Insightful		?		√

8 Future Work II — A Dynamic Decision-Making Experiment

8.1 The Experimental Task

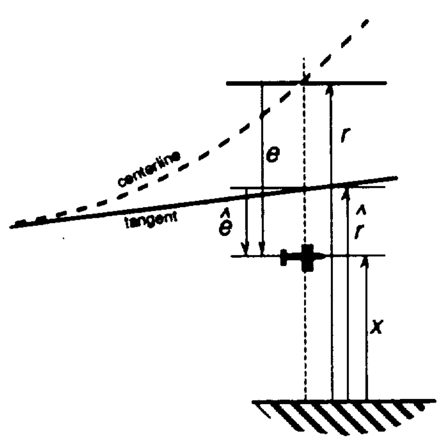
I am currently designing an attention-allocation experiment, in which a subject pilot will maneuver an aircraft along the center of a "lane". The subject will not, however, know his position relative to the center-line of the lane, only relative to the tangent to the centerline at a previous point in time — the last point of information acquisition. The task is shown below in Figure 10.

Figure 10: The experimental task: flying along a curved path. The pilot would like to fly along the dotted centerline, but only knows his position relative to the tangent to the centerline at the last point of "updating".



In this experiment, the pilot is essentially choosing between two competing tasks. He must either (i) pilot the vehicle manually along the current tangent (and risk being on a poor track which will lead outside the protected path), or (ii) update the tangent (and pay the penalty of having the compensatory display frozen for τ seconds while the tangent-track is updated). The relevant variables are depicted in Figure 11.

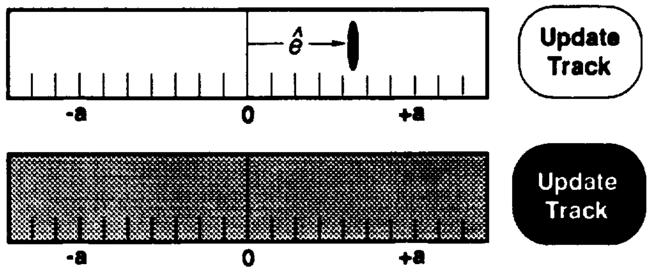
Figure 11: Control variables. x is the lateral displacement of the aircraft relative to some datum. r and $\hat{r}$ are the positions of the real target (the center-line) and the tangent, respectively, relative to the same datum. e and $\hat{e}$ are the aircraft's position errors relative to the center-line and the tangent respectively.



8.2 The Displays

The visual display, shown in Figure 12, is compensatory, as are many cockpit navigation displays (e.g. VOR and ILS). A rate display, such as a turn-coordinator, is also provided to aid the pilot in the secondary [manual control] task.

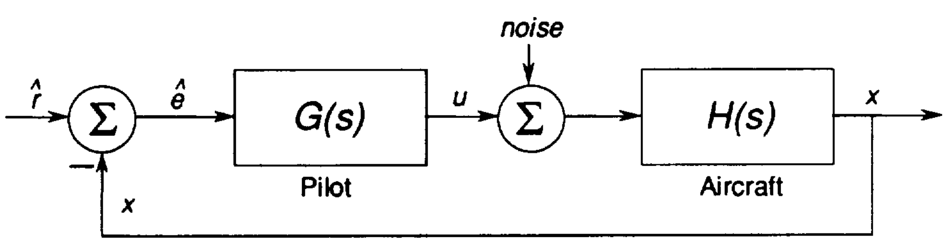
Figure 12: The visual display in its two modes (i) as a compensatory tracking display, and (ii) in the track update mode (note that the tracking display is temporarily grayed out).



8.3 The Vehicle Model

The vehicle is modelled as an unstable or marginally stable system, with noise added to the control input, as shown in Figure 13 below. In this way the subject's attention must be focused on the control task in order to achieve good control performance. The degree of instability and amount of noise will be adjusted to provide the required level of difficulty.

Figure 13: The vehicle dynamics



The aircraft transfer function, *H(s)* , currently modelled is

$$H(s) = \frac{x(s)}{u(s)} = \frac{K}{s(s-a)}. \qquad [5]$$

Since it must be simulated digitally on the PC, a backward-difference discrete approximation for the Laplace operator, s:

$$s \leftarrow \frac{z-1}{zT}, \qquad [6]$$

was used to determine the corresponding discrete transfer function between the control action and the vehicle state:

$$H(z) = \frac{x(z)}{u(z)} = \frac{KT^2 z}{z^2(1-aT) + z(-2+aT) + 1} \quad [7]$$

This transfer function is then coded as the following difference equation:

$$x_k = (2-aT)x_{k-1} + (-1+aT)x_{k-2} + T^2K.u_{k-2} \quad [8]$$

8.4 Scoring

At first, the costs will be made very explicit, so that we can test the optimality hypothesis. The subject's penalty for deviation from the curved path will be either: (i) a constant penalty, integrated over the time the aircraft is outside the lateral boundaries of the course:

$$J = C\int \delta_{outside}.dt \quad [9]$$

or (ii) a relatively severe penalty, accrued only very occasionally (probabilistically) when the vehicle is outside the boundaries of the lane.

Appendix A — Tables

Tables A1 through A7 show the percentage errors obtained when each model is applied to each data set.

Table A1: Percentage errors for the Utility model.

Model	Data D	Data H	Data P	Data R
D	29.0	47.2	42.4	43.3
H	32.3	13.9	9.1	10.0
P	19.4	25.0	6.1	3.3
R	12.9	16.7	21.2	13.3

Table A2: Percentage errors for the Perceptron, 4-D model.

Model	Data D	Data H	Data P	Data R
D	22.6	41.7	12.1	16.7
H	25.8	5.6	9.1	10.0
P	32.3	30.6	3.0	13.3
R	35.5	33.3	3.0	0.0

Table A3: Percentage errors for the Closed-form, 4-D model.

Model	Data D	Data H	Data P	Data R
D	41.9	47.2	30.3	40.0
H	45.2	33.3	27.3	30.0
P	45.2	33.3	18.2	23.3
R	45.2	33.3	18.2	23.3

Table A4: Percentage errors for the Closed-form, improved, 4-D model.

Model	Data D	H	P	R
D	41.9	38.9	36.4	43.3
H	45.2	33.3	30.3	30.0
P	38.7	25.0	18.2	23.3
R	48.4	33.3	18.2	23.3

Table A5: Percentage errors for the Perceptron, 2-D model.

Model	Data D	H	P	R
D	25.8	30.6	6.1	13.3
H	29.0	13.9	6.1	10.0
P	38.7	41.7	3.0	13.3
R	32.3	25.0	6.1	10.0

Table A6: Percentage errors for the Closed-form, 2-D model.

Model	Data D	H	P	R
D	48.4	36.1	21.2	26.7
H	48.4	36.1	21.2	26.7
P	48.4	36.1	21.2	26.7
R	48.4	36.1	21.2	26.7

Table A7: Percentage errors for the Closed-form, improved, 2-D model.

Model	Data D	H	P	R
D	35.5	30.6	15.2	20.0
H	38.7	30.6	12.1	16.7
P	45.2	36.1	18.2	20.0
R	41.9	36.1	18.2	20.0

Appendix B — Plots

Figure B1: Difference in Utility plotted against Subjective Difference for the four subjects

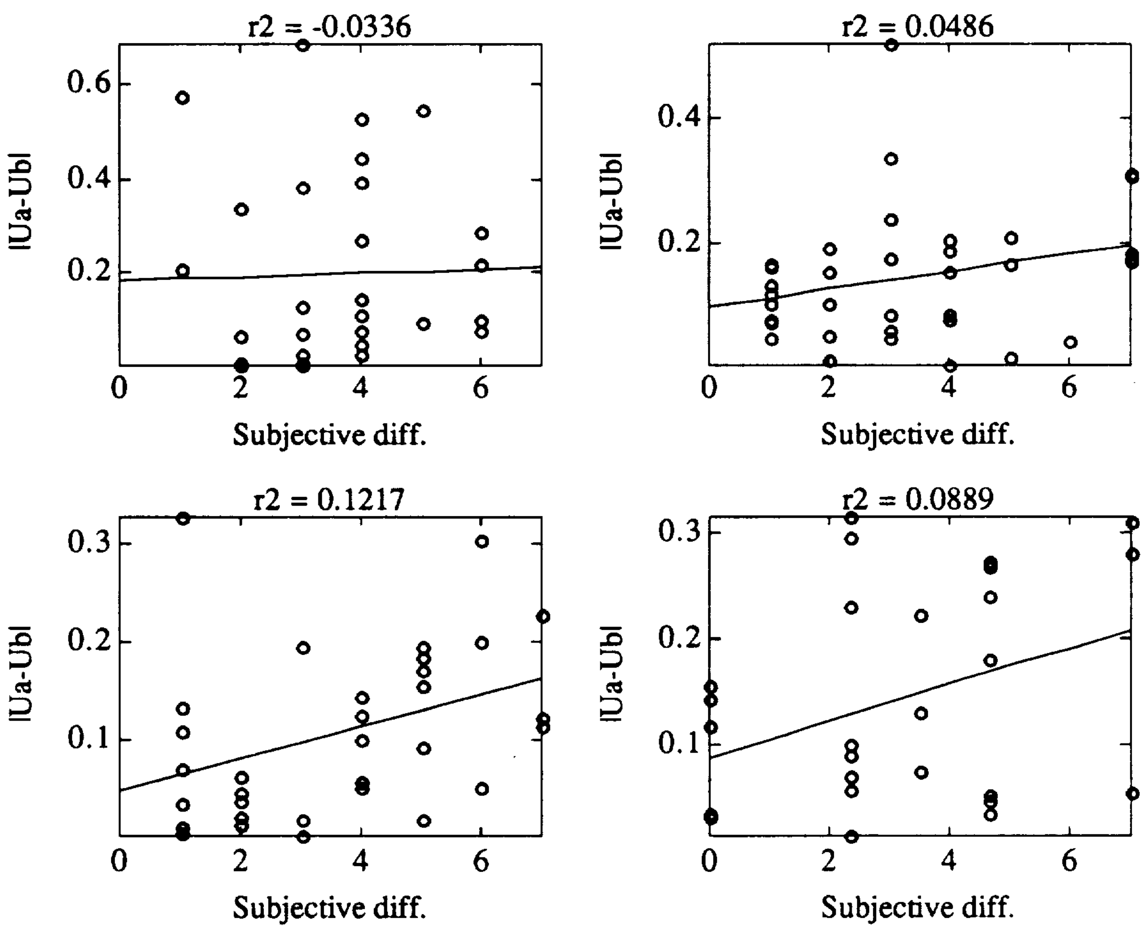


Figure B2: Distances from Decision Plane plotted against Subjective Difference in alternatives for the 4 subjects. Case: 4-D decision model, Perceptron solution. Correlation coefficients shown.

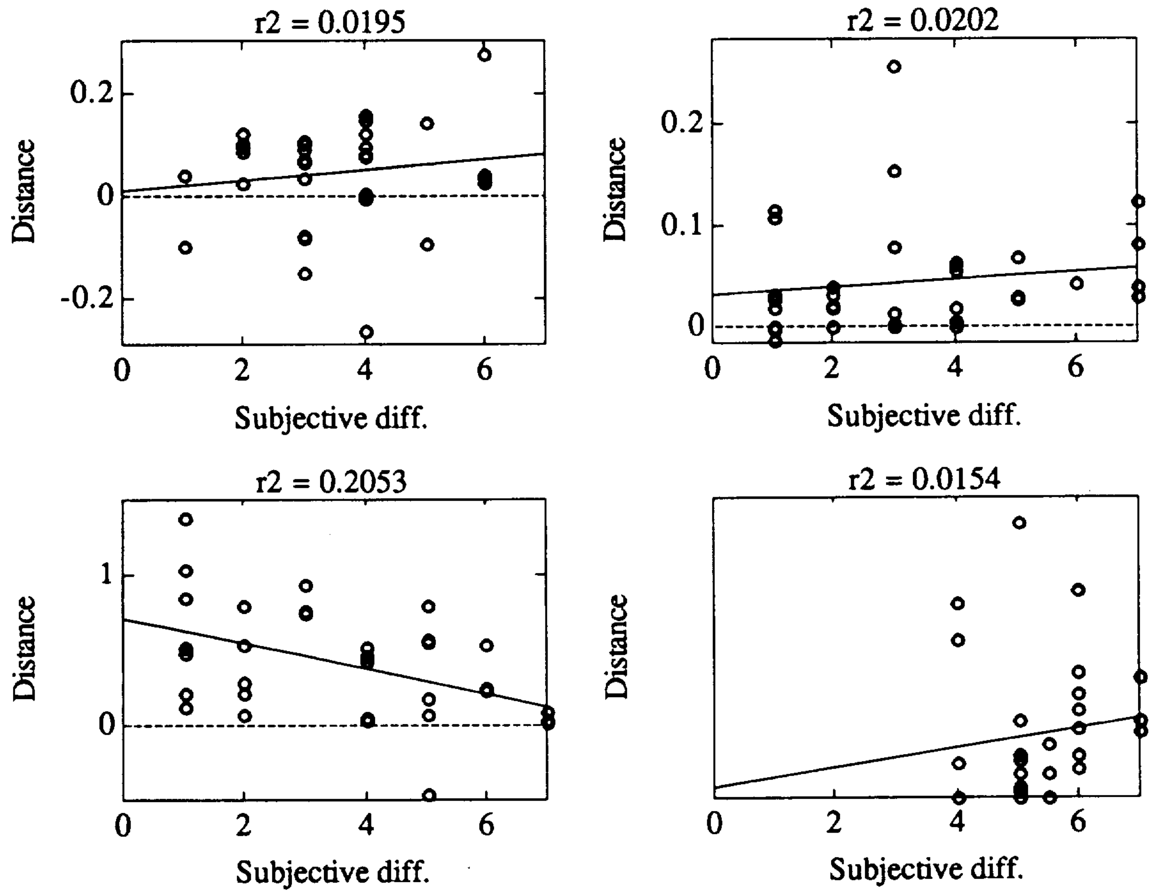


Figure B3: Distances from Decision Plane plotted against Subjective Difference in alternatives for the 4 subjects. Case: 4-D decision model, closed-form solution. Correlation coefficients shown.

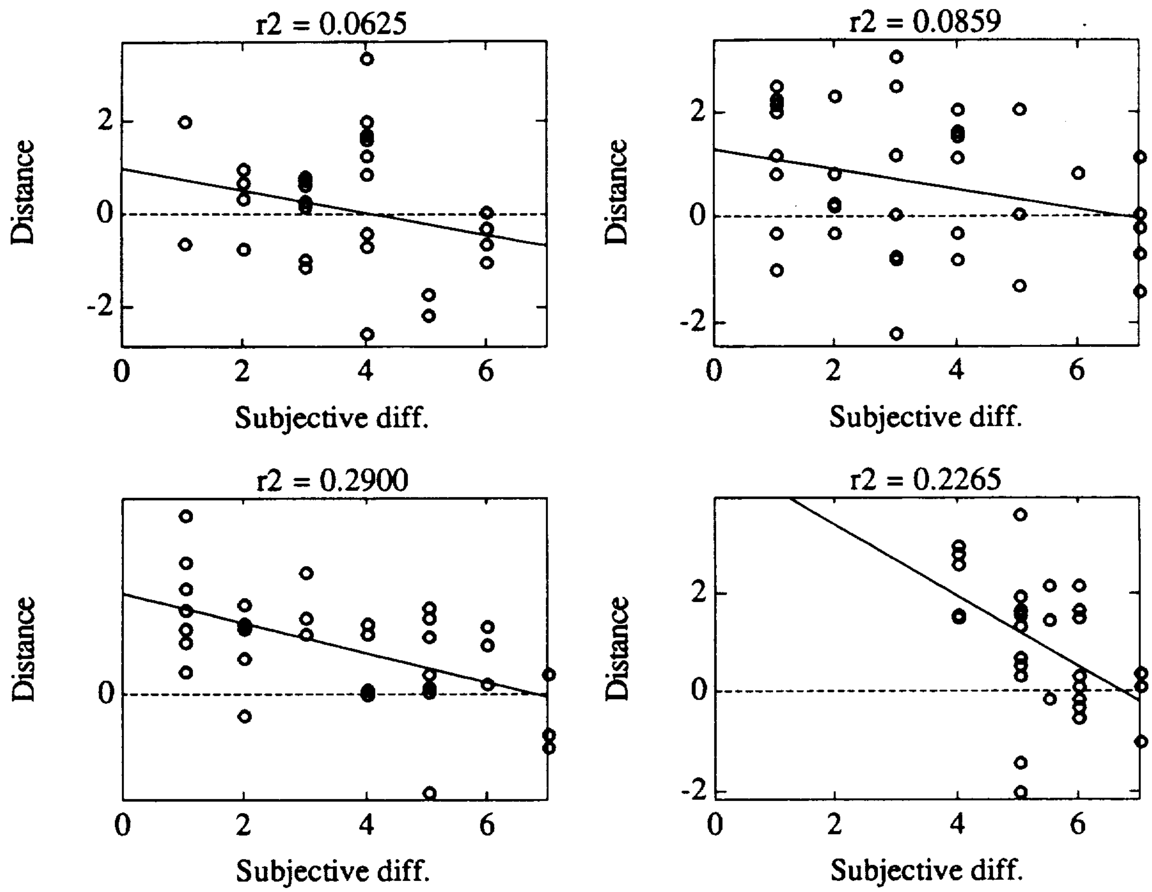


Figure B4: Distances from Decision Plane plotted against Subjective Difference in alternatives for the 4 subjects. Case: 4-D decision model, closed-form 2 solution. Correlation coefficients shown.

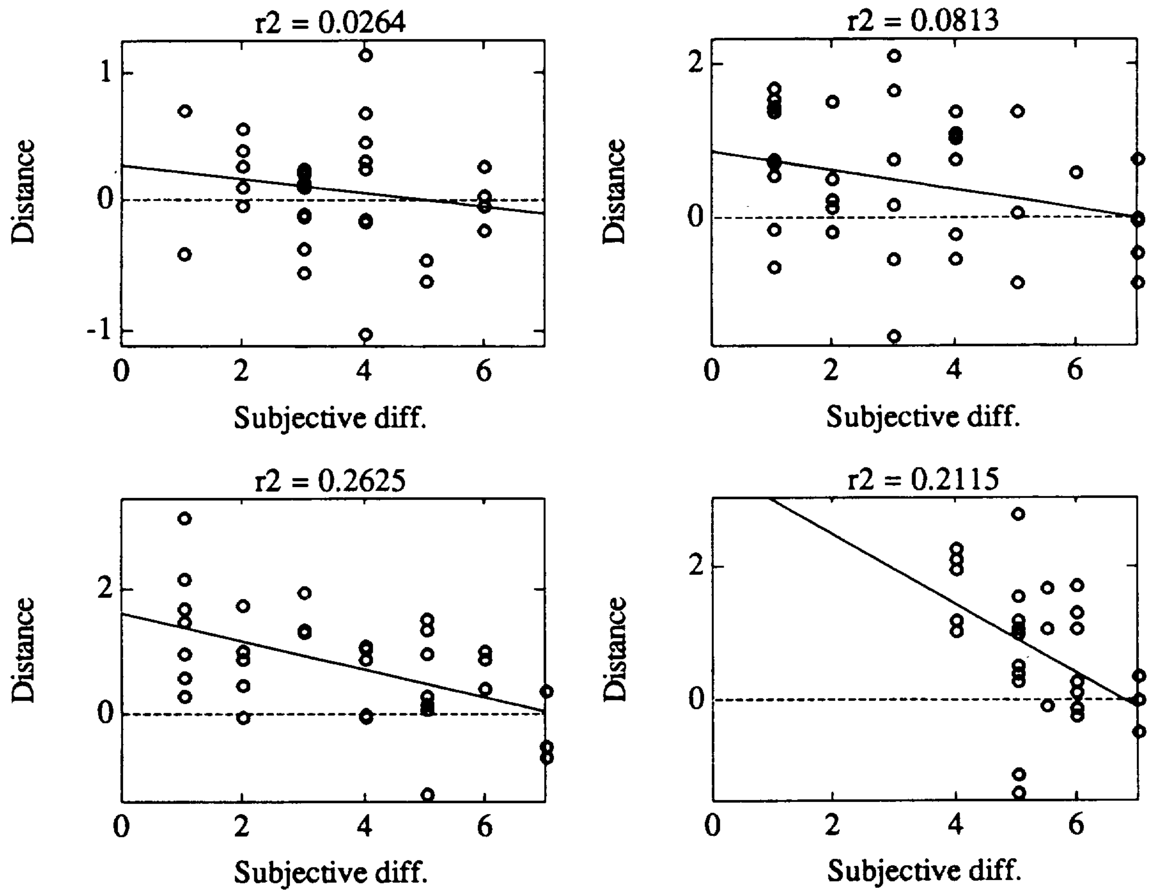


Figure B5: Distances from Decision Plane plotted against Subjective Difference in alternatives for the 4 subjects. Case: 2-D decision model, Perceptron solution. Correlation coefficients shown.

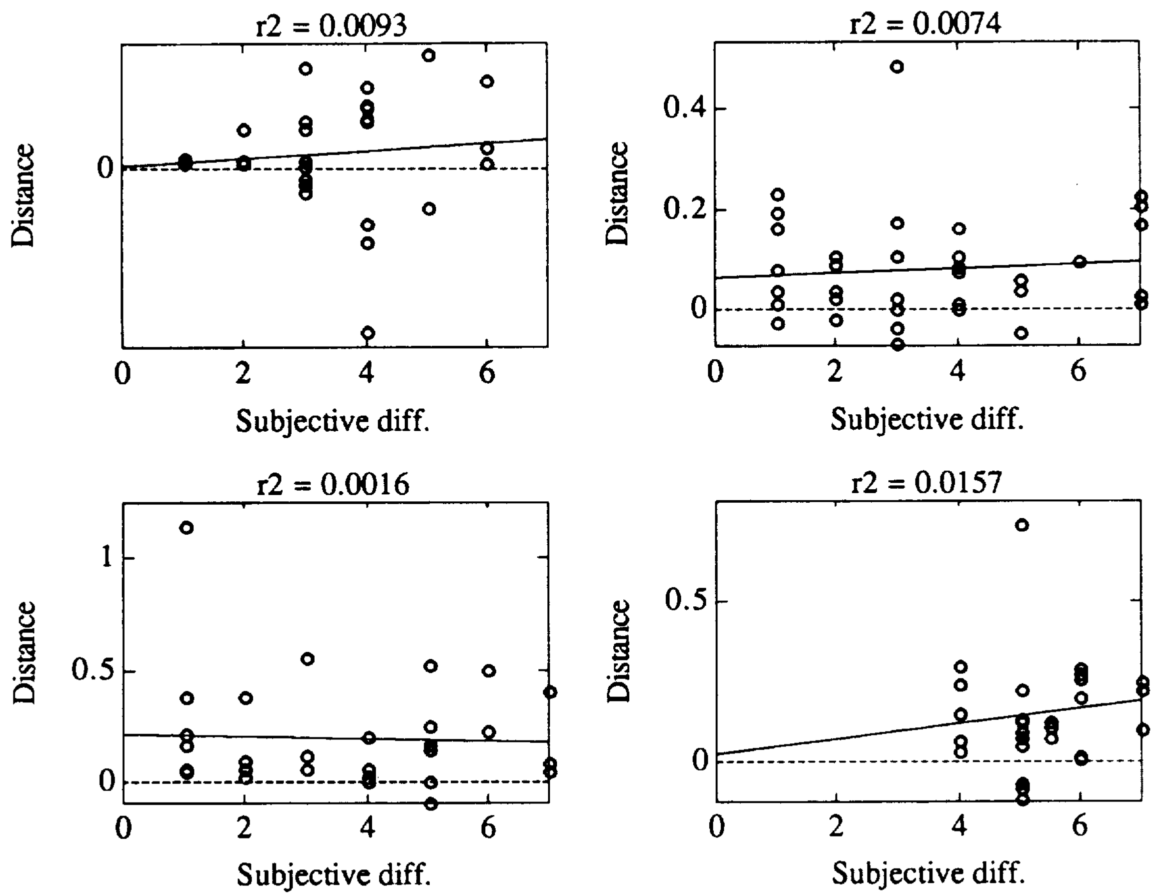


Figure B6: Distances from Decision Plane plotted against Subjective Difference in alternatives for the 4 subjects. Case: 2-D decision model, closed-form solution. Correlation coefficients shown.

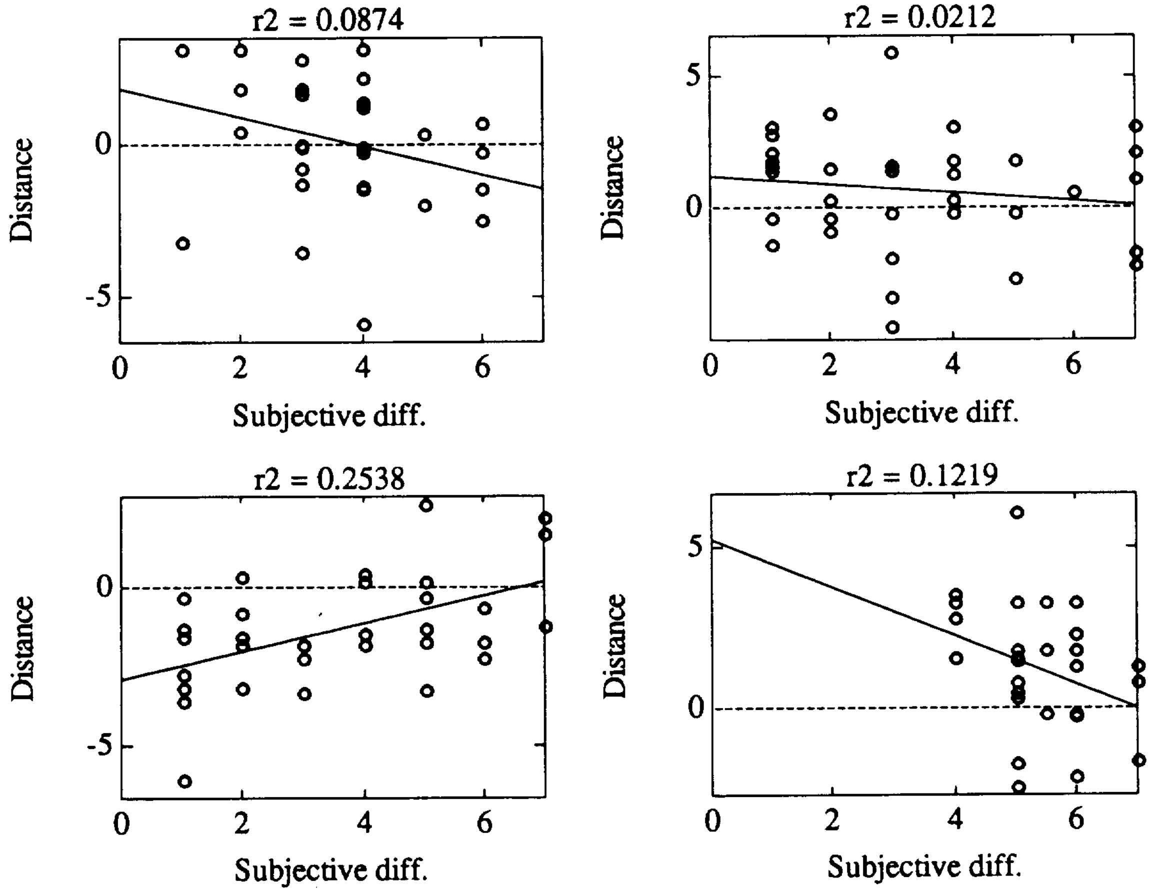


Figure B7: Distances from Decision Plane plotted against Subjective Difference in alternatives for the 4 subjects. Case: 2-D decision model, closed-form 2 solution. Correlation coefficients shown.