SA/PM 5

HOME OFFICE

SCIENTIFIC ADVISER'S BRANCH

POLICE MEMORANDUM

CRIMINAL INVESTIGATION FROM THE SMALL TOWN TO THE LARGE URBAN CONURBATION

by

M.A.P. WILLMER

This memorandum represents the personal views of the author. It does not carry the authority of the Home Office, Scientific Adviser's Branch.

MAY 1966

HOME OFFICE

SCIENTIFIC ADVISER'S BRANCH

Criminal Investigation from the Small Town to the Large Urban Conurbation

by

M.A.P. WILLMER

Summary

The paper looks at the problems which face criminal investigators as communities grow from small towns to urban conurbations. It begins by looking at the basic method of criminal investigation and goes on to show that the level of uncertainty associated with crimes only begins to decrease appreciably if there is a ten percentage probability or more that the set of suspects does in fact contain the criminal. The paper then makes the point that a better measure of C.I.D. success is the remaining uncertainty level rather than the detection rate. Next, using this measure, the success of the police for an area is estimated as it grows from a village to an urban conurbation. Included in this work are some results from a recent survey of crimes in a typical S.E. County Force. The difficulties which face detectives are then reviewed and the introduction of computers is suggested. Finally, the paper proposes some studies which should be undertaken to show how men and computers should be best integrated and how such a combination could be used most effectively.

Contents

List of Illustrations

1. Introduction

1.1. Over the past years criminal statistics have shown that the number of indictable offences have risen to a record high level. In 1964 there were in fact 1,067,963, such offences recorded for England and Wales. During this time the population has grown and the country is, in general, enjoying a record standard of living. Our towns and cities have also grown and in some cases it is difficult to know when one stops and another begins. In this way we have obtained large urban conurbations where people tend to work in one district, live in another and spent their money in a third. Opportunities for crime have increased and with more cars on better roads, criminals have been able to achieve a high degree of mobility. All these factors have made it more difficult for detectives to carry out their task of solving crimes because the information which they need is now harder to obtain. It is not surprising, therefore, that over the same period detection rates have fallen. Indeed in some parts of the country the detection rate for many classes of crime is extremely low e.g. for breaking offences it can be as low as twenty per cent.

1.2. From a scientific point of view the solving of a crime can be regarded as the gathering and use of information about people, things and events in an effort to associate a person or persons with a criminal act or acts with a high degree of probability i.e. beyond reasonable doubt. Failure, therefore, on the part of the police to solve a crime occurs because the necessary information is:

(a) not getting into the police system

(b) getting into the system but being lost or distorted

(c) getting into the system but not being used to the best advantage.

1.3. Many ways have been suggested in order to obtain an improvement in the situation. It is often suggested, for instance, that what is needed is more detectives or that they should operate in a different way than at present. The formation of regional crime squads is an example of reorganising detectives. Other suggestions are that the police should have better equipment or that they should be able to see the state of people's bank accounts. Essentially, the suggestions have three factors in common:

(a) men and the way they are used

(b) equipment

(c) the rules under which the police have to operate.

1.4. What appears to be lacking is a concept of criminal investigation which shows how these factors are interconnected and interrelated and which can be used to estimate what improvements in manpower and equipment, methods of operation as well as what changes in the restrictions imposed upon police activities are necessary in order to achieve a given increase in the level of detection success. It will then be for society to decide whether it is prepared to pay the price in terms of extra money or less personal freedom.

1.5. In this paper I propose to show that this gap can be filled if the investigation of crimes is considered as an attempt to reduce the level of uncertainty or entropy connecting the people of a society and the crimes connected within it. In ref. 1, I showed that the value of information could be measured, in principle, in terms of the amount it changed the uncertainty or entropy level. In this reference, I concentrated mainly upon a single crime and it is now necessary to think in terms of the effect of information on the uncertainty or entropy levels of many crimes. In this way we can think of the total level of uncertainty for a police district changing as the police system receives information. This outstanding uncertainty or entropy is a

measure of the work still to be done. If, therefore, we take the outstanding entropy of the crimes when the police gave up, either because further progress was physically impossible or because of a lack of manpower or equipment to continue the investigations, we have a measure of the failure of the police. Because such a measure takes into account the partial success of the police it is fairer and more accurate than failure ratio or its complement, the detection rate.

1.6. It is also fairer to the police in another respect. In any society where the freedom of the individual is highly treasured the actions of the police will always be restricted. This means that in some cases valuable information will be prevented from reaching the police system. In some such cases, it is to be expected that the police will be unable to obtain information of comparable value within the restrictions placed upon them. In these cases there will exist an entropy residue which the police are powerless to reduce. Hence, it follows that a truer measure of the success of the police in their efforts to solve crime is given by the difference between the outstanding entropy after the police have completed their investigations and the residual entropy which can be attributed to the restrictions on their activities.

1.7. I will start by examining the basic method of criminal investigation used by the police today. After discussing, in broad terms, the type of information which is required, I will look at the way the changing nature of society from small town to urban conurbation could alter the value of different pieces of information as well as the frequency and probability that they are received by the police network. The problems and advantages of integrating modern automatic data processing equipment into the police network are then discussed. Finally I suggest a possible programme of study to determine:

(a) how such an integration could best be made

(b) how to use the combination most effectively

(c) the effect of the present restrictions on police activities.

2. The Basic Method of Criminal Investigation

2.1. Fig.1 shows diagrammatically the way in which the problem of solving crimes is tackled. Although the diagram shows only two crimes, it can be extended, in principle, to any number.

2.2. When a crime is reported, the police go to the scene to investigate. From their searches and enquiries at and in the neighbourhood of the scene, they obtain information which conveniently falls into two basic categories. Firstly there is active information which is information that leads to the establishment of a set of suspects. The strength of the suspicion against each member of this set will be dependent upon the nature of the evidence accumulated. Secondly there is passive information which is information that can only be of use if a set of suspects has been obtained. By the use of this passive information the suspicion against each member of the suspect set i.e. the probability that each member of the suspect set can be associated with the crime, can be either increased or decreased. These two basic types of information lead by two routes, see fig.1, to the "final state of uncertainty". In those cases where there is sufficient evidence the appropriate person or persons are, in general, charged and sent for trial.

2.3. Information about people, things and events also reach the C.I.D. from the public, informants, beat patrol policemen, their own internal local knowledge etc. Such information can either produce a number of suspects for a given crime or else help establish a pool of general suspects i.e. people suspected of being actively engaged in crime. This information can be combined with [illegible] the scenes of crimes in order to influence the final state of uncertainty.

2.4. From the above it would appear that the establishment of a set of suspects plays a vital role in the solution of crimes. Let us use the concept of entropy to examine this point further. We shall assume that our investigations have led us to the stage where there are n persons in the suspect set all of whom are equally suspect. It can be shown, following ref. 1, that, if the probability of the criminal being in the suspect set is p, the entropy level associated with the crime is given by

$$H = -p \log_{10} \left(\frac{p}{n}\right) - (1-p) \log_{10} \frac{(1-p)}{(N-n)} \quad \text{pui} \qquad (1)$$

where N is the size of the total eligible population. This entropy level is plotted against the probability of the criminal being in the suspect set in fig.2 for various values of n. It can be seen that unless p is greater than 10% the number of persons in the suspect group has little or no effect on the entropy.

2.5. The above analysis shows how essential it is for the police to be able to form a suspect set with a high probability that the set contains the criminal. In practice, the police attempt to form such a set from a combination of the information obtained from enquiries, their local knowledge, records, informants and other forces etc. In some cases this information will lead to a small suspect group whilst in other cases, especially for the smaller and more trivial crimes, the size of the possible suspect set will be very large. Yet even in some of these cases the police have pieces of passive information which could on some occasions be very useful if only the size of the suspect set were of manageable proportions. Another way of looking at passive information is to consider it as information which is associated with a person and which is either not recorded or the means of retrieving it from the records are not available. An example of the first type is given by the present police record system which only contains data about criminals. Fingerprints are a good example of the latter type in so far as the police have the fingerprints of criminals but only have the means to identify an S.O.C. print by searching the records if it belongs to a particular type of criminal. The information which the police have about the non-criminal section of society, in general, is either stored in the minds of individual police officers or is of a general nature concerning a group of the non-criminal part of the population not particular individuals.

2.6. With the decrease in detection rates, the growth of the population, the increasing ownership of the motor car and fast roads etc. policemen do not know so much about the people both criminal and non-criminal as they did years ago when populations were more self-contained and isolated. The effect of these changes in our way of life will affect the value of pieces of information which the police receive. The following analysis demonstrates this effect.

2.7. Let us suppose that a piece of information is received at the scene of a crime which could link a person with the crime, e.g. the criminal had red hair, was left handed, smoked an uncommon brand of tobacco, etc. Let us suppose further that the society contains N people n of whom on average could be expected to satisfy the conditions required by the piece of information. We shall assume that a proportion of the population, λ, belong to the group of the population of whom the police have records and the means to interrogate them. In this case we can show that the value of the item of information, assuming that, before searching the records available, all people are equally likely is

$$\Delta H = \lambda \log_{10} \left(\frac{N}{n}\right). \qquad (2)$$

which is plotted in fig. 3 for various values of λ and $\frac{n}{N}$. In a small village when the police know everyone we would expect λ to be approximately equal to unity. In a large city this value could well fall to 20 - 30% and thus the value of the same piece of information will fall accordingly.

2.8. So far we have considered only a single crime. However, in practice crimes are often related and information which the police receive can on many occasions influence the entropy level of a number of crimes. Because, therefore, our changing pattern of living affects the final entropy level distribution of all the crimes in a given police district we shall consider the growth of a small village community to a town and thence to part of an urban conurbation.

3. From Village to Urban Conurbation

3.1. Let us consider a district in which, during a given time interval, the police were able to establish a final probability set for the nth crime of $\{ p_i^{(n)} \}$ where $p_i^{(n)}$ = the probability that the i^{th} person of the community committed the n^{th} crime.

From this probability set we can calculate the final entropy level $H^{(n)}$ where

$$H^{(n)} = -\sum_i p_i^{(n)} \log p_i^{(n)} \quad \text{pui} \qquad (3)$$

where the pui is the measure of entropy, see ref.1.

From the entropy level obtained from the N crimes we can construct a histogram from which we can estimate the probability distribution associated with the final entropy level. Thus, if $p_H dH$ is the probability that the final entropy level lies between H, and H & dH, the remaining uncertainty which the police have been unable to reduce during the course of their enquiries is given by

$$H_R = N \int p_H H \, dH \quad \text{pui} \qquad (4)$$

The value of $\frac{H_R}{N}$ can be used as a measure of the success of the police efforts to solve the crimes. A high value of H_R means lack of success and vice versa. It is a more accurate measure of police performance than pure detection rate because it includes those cases where the level of uncertainty has been lower but the criminal has not been caught.

3.2. Let us now consider a small isolated village community of, say, 200 people. In such a community the police will know all the inhabitants as individuals and there will be little that goes on that does not reach them. For instance, anyone who steals is likely to be suspected when the proceeds of their theft are brought into use. For such a district, a typical distribution of w(M), the probability that, at the end of police investigations, there are M equally likely suspects is given by

M	1	5	50	100	200
w(M)	.80	0.05	0.07	0.03	0.05

In this case the value of h_R the expected remaining entropy per crime is given by

$$h_R = \sum_M w(M) \log_{10} M \quad \text{pui per crime} \qquad (5)$$

giving $h_R = 0.33$. pui per crime

3.3. When, however, the village grows and develops into a small town it becomes more difficult for the police to know everyone and to have the same knowledge of what is going on in the district. In this we would expect a typical distribution of w(M) to be given by:

M	1	5	100	500	1000	5000
w(M)	0.70	0.06	0.05	0.04	0.05	0.10

giving an expected remaining entropy level per crime of 0.77 pui.

3.4. Let us now consider a set of towns which are close together and where movement from one to another is both quick and convenient and where many people work in a nearby city and commute daily. Now the size of the non-individual set is large, especially if the area contains a large number of newcomers, such an area is typical of the South Coast Counties today. From a recent Home Office questionnaire details were obtained about all the crimes of the housebreaking and burglary type from most police forces in England which:

(a) had been detected or written off during the week commencing the 6th December 1964

and (b) undetected crimes which were reported or discovered during the week commencing the 11th October 1964.

The results from a county force in the S.E. of England were analysed and after making several broad assumptions with regard to the probabilities involved, the final entropy levels were calculated. These results are given in fig. 4 where a histogram of the frequency with which each entropy level occurs is shown. From this data an expected remaining entropy level was calculated giving h_R = 1.88 pui per crime.

3.5. Fig. 4 shows that, to a large extent, crimes are either solved or are 'brickwalls' i.e. the police have no idea who commit them. It is interesting to note that for areas which are similar to the one under consideration, the expected remaining entropy per crime is approximately proportional to the detection failure ratio. This follows because of the very small percentage of crimes in which the police were able to reduce the uncertainty but unable to solve the crime.

3.6. In order to understand the reasons why police performance should decrease as communities grow let us consider further the problem of how information is received, interpreted and transmitted in the police network.

4. The Reception, Interpretation and Transmission of Information

4.1. The elementary functions of a policeman are basically to receive information, to interpret or process it and to respond to it. Fig. 5 shows the situation in diagrammatic form giving a few examples of the possible actions which can be selected. Now the effectiveness of the interpreter unit will, to a large extent, be dependent upon the size of its memory and the usefulness of its contents. Thus a policeman with a more effective interpreter unit will be better able to assess the value of information which he receives than a policeman with a less effective one. Indeed, a person or machine with all the information concerning a society and its people would be able to assess the value of a given piece of information most accurately.

4.2. It is possible, therefore, to imagine information flowing into the interpreter unit of a policeman, see fig. 5 each piece having a given value as determined by the machine with all the information. Because of the policeman's limited memory and lack of knowledge he will interpret the information he receives differently and thus the output to the action selector unit will be different from the input. Theoretically we can define the value of each piece of information received in terms of its pui rating using the concept of potential entropy outlined in ref.1. It is to be noted that a machine with all the information can make the most accurate estimate of the probability and entropy involved in contrast to the ordinary policeman. Fig. 5 again shows diagrammatically a typical input and output of information, and it can be seen that the limitations of the human mind make the interpreter unit behave in the same way as an interference generator in more typical information flow problems.

4.3. The performance of each policeman's interpreter unit will be dependent upon his training, local knowledge, the type of locality in which he is working, etc. If he fails to detect a potentially valuable piece of information, he may well neglect to perform a vital action. Also the performance of the system, i.e. a police force, will depend upon the judgements of its individuals. It is necessary, therefore, for each policeman to be able to detect information which is of

importance to one of his colleagues, and to transmit it to him as faithfully as possible. Failure to do so will lead to the system giving a non-optimum performance.

4.4. To illustrate how the effectiveness of the interpreter unit can affect the performance of the policeman as part of a communication network let us consider a simple example. Because the interpreter unit receives and transmits information which can be measured in terms of pui, the situation is similar to that of general communication theory where the effect of interference in the system is measured by means of its effect upon the rate of actual transmission, R, defined by

R = Entropy of the input - the entropy of the input when the output is known see ref. 2. This equation represents the amount of information transmitted corrected to take into account the amount of this information which is missing in the received signal. Thus, if we assume that the input information can be divided into two groups viz. a group with a high potential pui value and a group with a low one where

P = probability that a piece of high pui value information is received
Q = probability that a piece of low pui value information is received
p = probability that a piece of high pui information is interpreted correctly
q = probability that a piece of high pui information is interpreted wrongly

plus the assumption that a piece of low pui information is always interpreted correctly, see fig. 6 it can be shown that the rate of actual transmission for such a policeman is given by

$$R = -\left(P \log_2 P + Q \log_2 Q \right) + P_q \log_2 \left(\frac{P_q}{Q+P_q} \right) + Q \log_2 \left(\frac{Q}{Q+P_q} \right) \qquad (6)$$

This function is plotted against q in fig. 7 for various values of P and Q. From this figure it can be seen that the actual rate of transmission which represents the success of a policeman as a receiver of information gathered during the course of his work is closely dependent upon the probability that the interpreter unit will miss a piece of high pui information.

4.5. So far in this section we have only considered the case of an individual policeman. In practice, however, the police network is composed of a number of policemen backed up by library facilities such as C.R.O.s which can be interrogated by policemen or their civilian associates. The flow of information in this case between the outside world and the police network is shown diagrammatically in fig. 8a., where, for the sake of simplicity, the police force is assumed to consist of three people, one of whom looks after the records. The diagram can easily be generalized for a force of any size.

4.6. As stated earlier in this paper, the object of this flow of information is to decrease the level of uncertainty or entropy which is associated with the crimes committed in the police district under consideration. The network is most efficient when each member is able to interpret the information which reaches him, to be in the position to collect the maximum amount of useful information and to select the best action following the receipt of such information. For a small community like a small town, we can expect such a system to be reasonably satisfactory because it is possible for man's interpreter unit to be able to cope successfully with the information that reaches him. The trouble arises when the size of the community develops and becomes a vast urban conurbation so that it is impossible for each policeman to receive, retain and use all the information which he needs to interpret successfully the incoming information.

4.7. At the present time, the police service attempts to deal with this problem by training men to make the best interpretations and decisions possible under the circumstances. As far as the author is aware, no use is made of modern automatic data processing equipment which can have vast memory stores, make interpretations

and select destinations for the interpreted information in order to back up the operational policeman. A possible network incorporating such equipment is shown diagrammatically in fig.8b. In such a system each man can communicate with the computer directly.

4.8. If such equipment were to be introduced into the system, it would mean training men to use a new set of decision rules as far as incoming information is concerned. It is to be noted that the formation of intelligence sections is a first step in this direction using men as the computer. The introduction of modern processing equipment has the advantage that it could cope with very much more data than is the case at the moment. In such a system the object of the computer would be to store and interpret information and then select the person or persons to whom this interpreted information should be sent. In this way, it is theoretically possible for men to receive information from the computer which they would recognize as having a higher pui value than that which would have been given to it if the interpreting was only done by the men of the system themselves. In practice, however, there will be many occasions when a man must decide what action to take immediately after receiving it so that here a computer will be of no value. Further, it will be physically impossible for a computer to receive all the information which the police system receives so that decisions will have to be made regarding what information should be sent to the computer. Also, it is impossible for all police experience to be recorded in the computer so that its power of interpretation will be limited and it may well be that, on many occasions, the policeman's interpretation will be the better. However, the computer is a tool which can handle vast quantities of data so that it has a potential which would be useful to the police in their efforts to bring criminals to justice. Because of this potential it is necessary to determine how men and computers can best be integrated. In other words, we must determine how computers can best be used in order to be of greatest service to the police force.

5. Towards the Integration of Men and Computers

5.1. In order to be in a position to determine which is the best way to integrate men and computers as well as to ensure that the combination is being used most effectively, it will be necessary to obtain data about all the various aspects of the problem.

5.2. Firstly, we should study the input of information into the system. Here it will be necessary to ensure that each man who collects information from the outside world is used to the greatest advantage. Some examples of studies which could be valuable in this connection are:-

(a) how to allocate beat patrol-men to ensure that the chances of meeting housebreakers or possible house-breakers in transit are maximised;

(b) how to reduce the time taken by the police to respond to emergency calls;

(c) how to ensure that beat patrol-men are patrolling in the most effective way, i.e. walking, bicycle, motor cycle or car;

(d) how to allocate resources most effectively, e.g. should operational statistics be used as an aid to force deployment?

(e) the use of network techniques to increase efficiency;

(f) the use of informers and whether it is profitable to increase their number;

(g) how to make policemen more crime minded so that they will become more aware of events about them.

5.3. Such studies as those listed above are aimed at trying to ensure that as much useful information as possible has a chance of getting into the police system. The next important problem which has to be studied is how to ensure that it is used

effectively once it is inside. To achieve this goal, it is necessary that such information is interpreted correctly. This raises the question of how policemen should be trained, i.e. how should they be taught to interpret information. A possible start to this problem is to examine how the intrinsic value of the many types of information which the police can receive have changed due to changes in the nature of our society and environment. In other words what is being lost by the development of towns into urban conurbations.

5.4. The next important problem which needs to be studied is that of how men should be trained to decice upon which course of action they should take after receiving each item of information. It is to be expected that with the introduction of a computer into the system a different set of 'rules' will have to be adopted than those currently used. Perhaps the major question that will have to be answered is when should information be sent to the computer.

5.5 Very closely allied to this problem will be that of what specific jobs should a computer be programmed to tackle. In principle, it will be to store and interpret information as well as select those people to whom the information should be sent. Its function is basically to ensure that men receive information which they can recognise as having a higher pui value than would be the case without a computer. In practice, its usefulness will probably be dependent upon the ability of scientists and programmers to find suitable tasks for it to do as well as on the magnitude of the probable cost. On the type of task which the computer is programmed to do will depend the type of information required from policemen. Some examples of the sort of work which could be given to a computer are:-

(a) to analyse the contents of very large crimes where there are large numbers of statements and a great deal of detail in order to look for inconsistencies;

(b) to combine operational statistics with records, sociological and criminological data to suggest suspects and predict future crimes;

(c) to computerize the work of intelligence sections - not to replace men but to increase their power and scope;

(d) to advise policemen of the indentity of any suspicious persons which they come into contact with during their work;

(e) to advise men on whether any car which they see is stolen;

(f) to compare S.O.C. marks with the large fingerprint reference collection.

Only studies and experiments will show whether such ideas are worth the large expenditure which would be required to put them into general operational use.

5.6. Lastly, it is important to remember that despite the introduction of modern equipment or increasing the size of forces, the success of criminal investigation will probably always be affected to some extent by the restrictions placed upon the police by society. It may well be, therefore, that if society demands a given level of success from its police force some of these restrictions will have to be relaxed. The effect of relaxing some of the restrictions would allow the police to gain more information about a wider range of people i.e. to increase the value of λ see para. 2.7. Fig. 3 shows that such an increase would lead to an improvement in the value of some types of information which the police can obtain. It is important that this aspect should be studied and the probable effects of easing some of the less important restrictions estimated.

May 1966

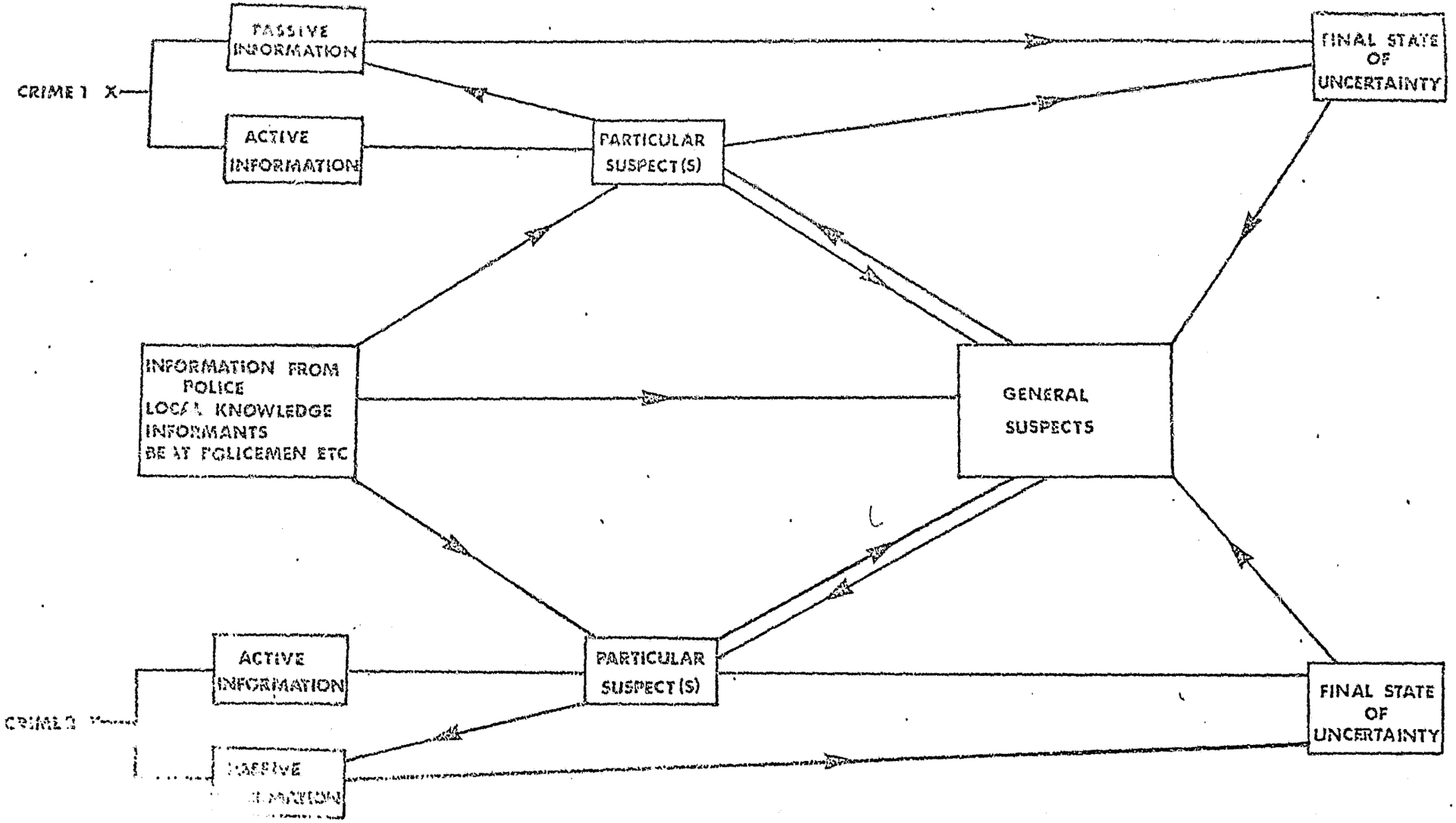
References

1. WILLMER, M.A.P. On the Measurement of Information in Fields where the Meaning of Messages is Important.

Operational Research Society Conference
Shrivenham, Wilts. September 1965

2. SHANNON, C.E. and WEAVER, W. The Mathematical Theory of Communication

University of Illinois Press Urbana 1947.



J. SLOWIN. THE IMPORTANCE OF SUSPECTS IN CRIMINAL INVESTIGATIONS

FIG. 1

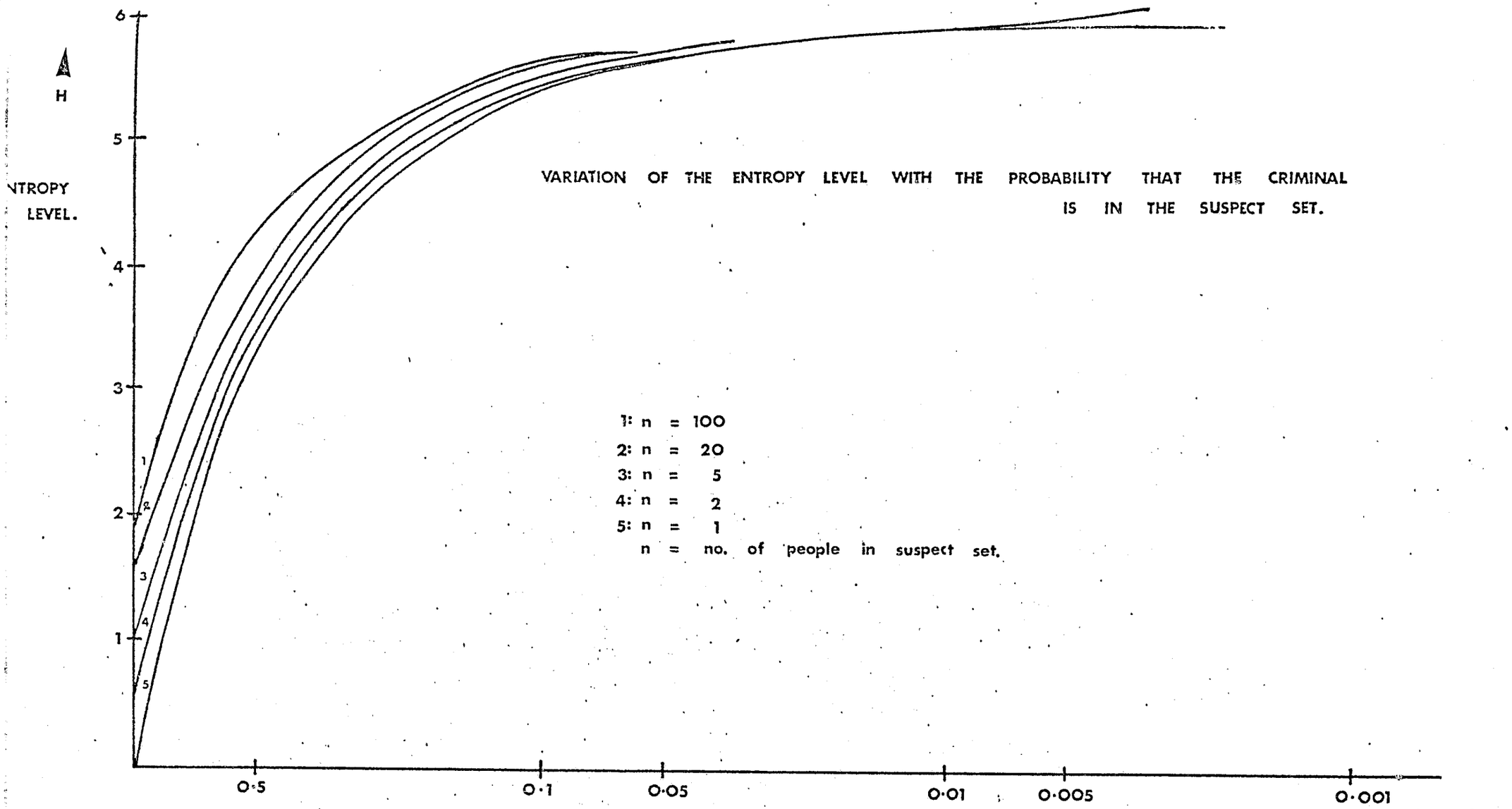
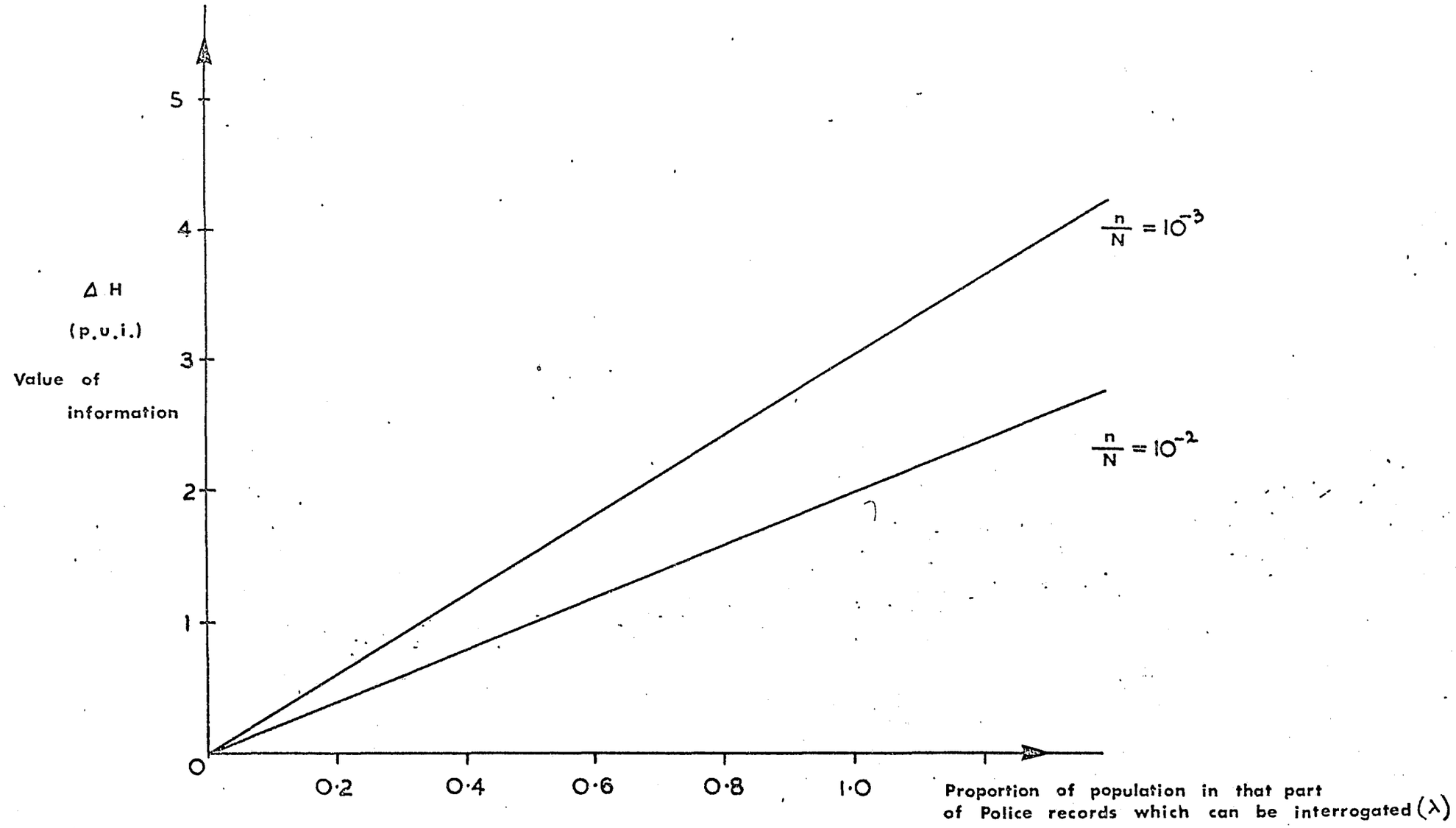
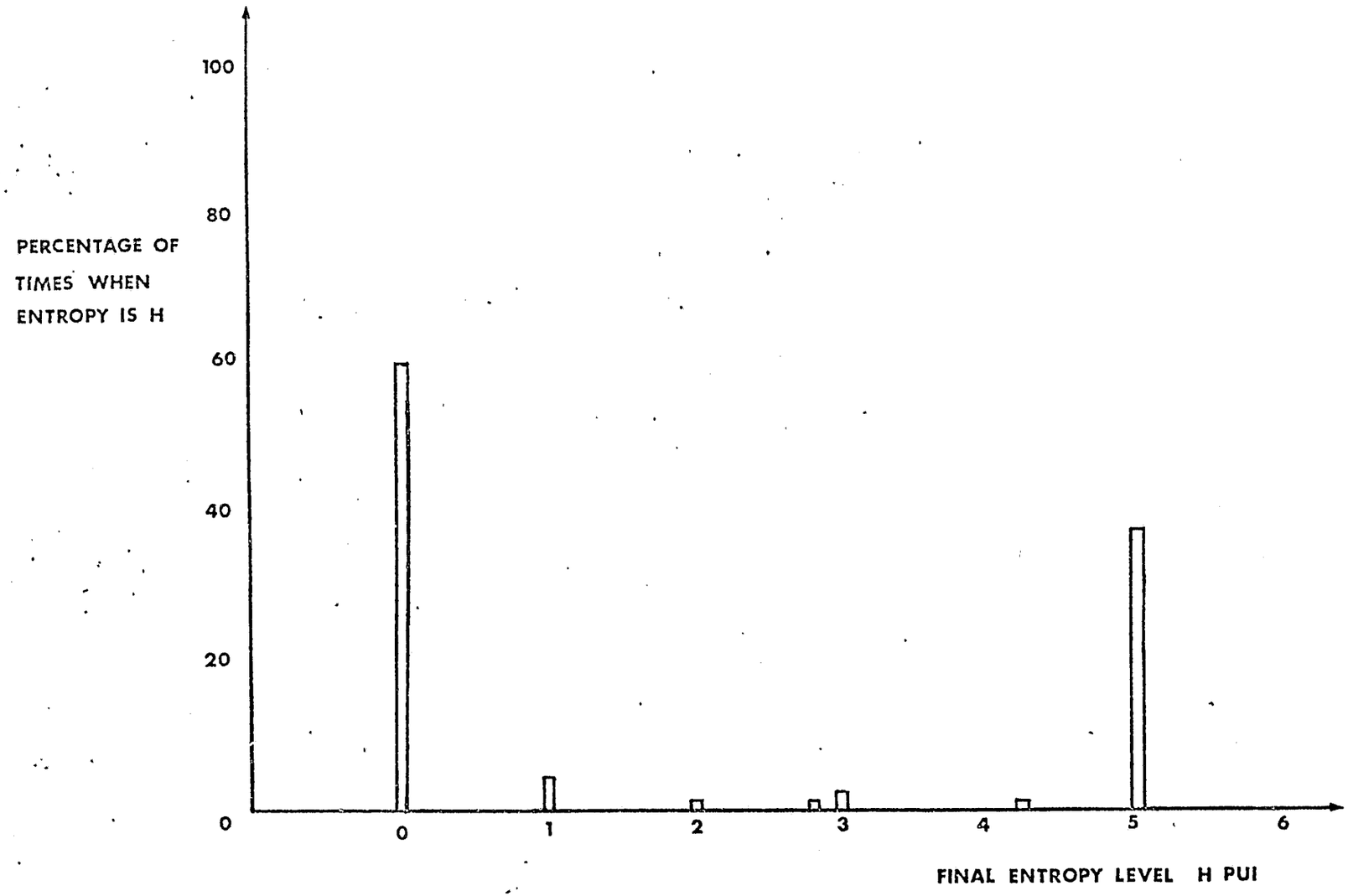

H
NTROPY
LEVEL.
6
5
4
3
2
1
VARIATION OF THE ENTROPY LEVEL WITH THE PROBABILITY THAT THE CRIMINAL IS IN THE SUSPECT SET.
1: n = 100
2: n = 20
3: n = 5
4: n = 2
5: n = 1
n = no. of people in suspect set.
0·5
0·1
0·05
0·01
0·005
0·001
P
PROBABILITY THAT SUSPECT GROUP CONTAINS CRIMINAL.

FIG. 2.



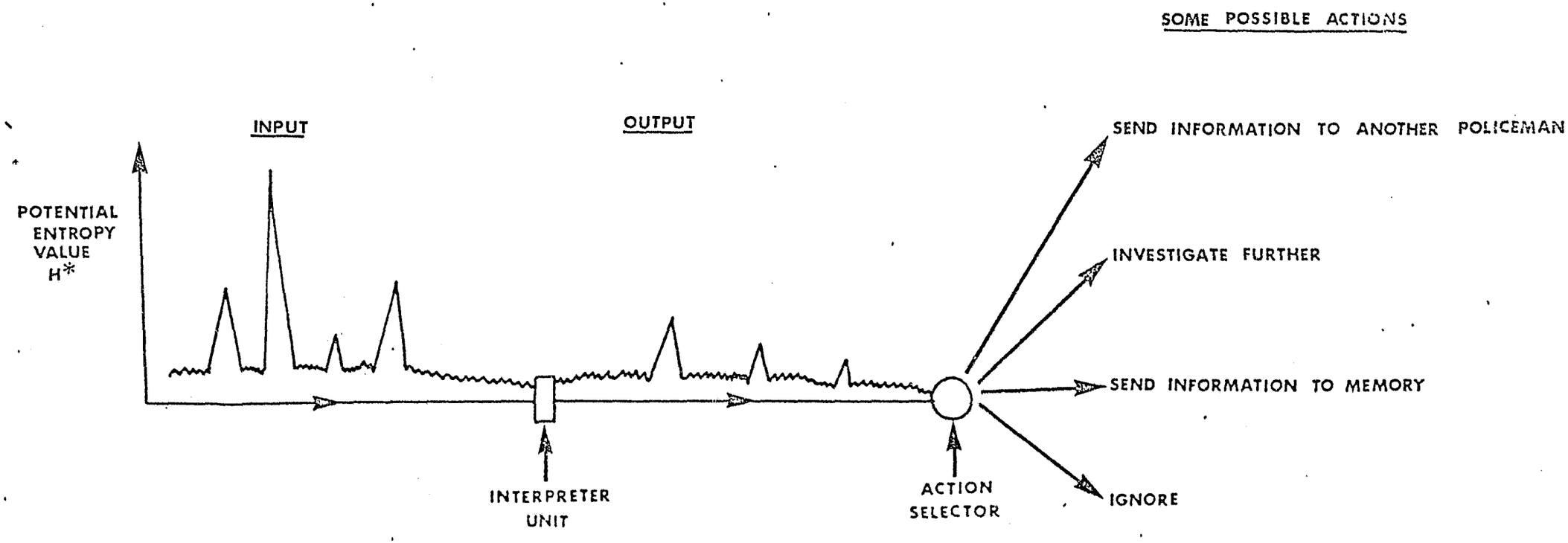
THE VARIATION OF THE VALUE OF A PIECE OF INFORMATION WITH THE PROPORTION OF THE POPULATION ON WHOM INFORMATION IS KEPT IN POLICE RECORDS.

FIG. 3



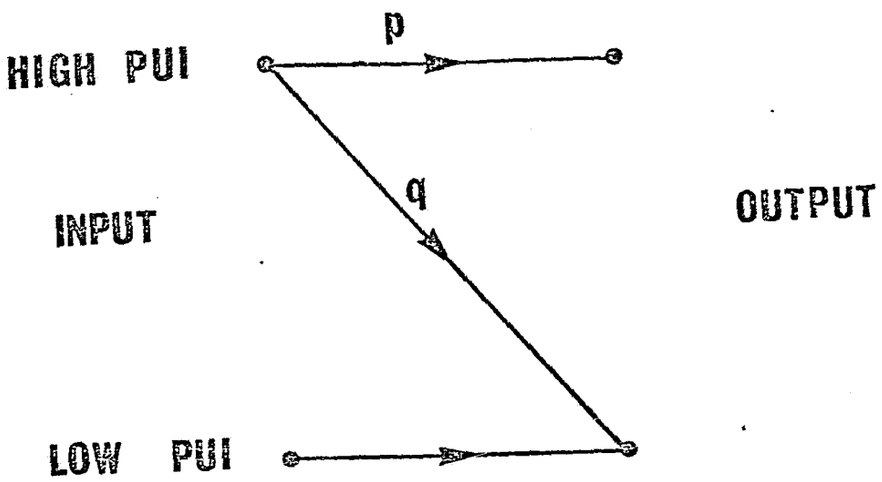
HISTOGRAM OF FINAL ENTROPY LEVELS FOR A TYPICAL COUNTY FORCE IN THE SOUTH EAST OF ENGLAND

FIG. 4



DIAGRAMMATIC REPRESENTATION OF THE INTERPRETATION OF AND THE RESPONSE TO INFORMATION RECEIVED BY A POLICEMAN

FIG. 5



P = Probability that a piece of high pui information is received

Q = Probability that a piece of low pui information is received

p = Probability that a piece of high pui information is interpreted correctly

q = Probability that a piece of high pui information is interpreted wrongly

DIAGRAM SHOWING HOW INPUT CAN BE WRONGLY INTERPRETED

FIG. 6.

VARIATION OF THE ACTUAL RATE OF TRANSMISSION WITH THE PROBABILITY THAT A PIECE OF HIGH PUI INFORMATION IS MISINTERPRETED.

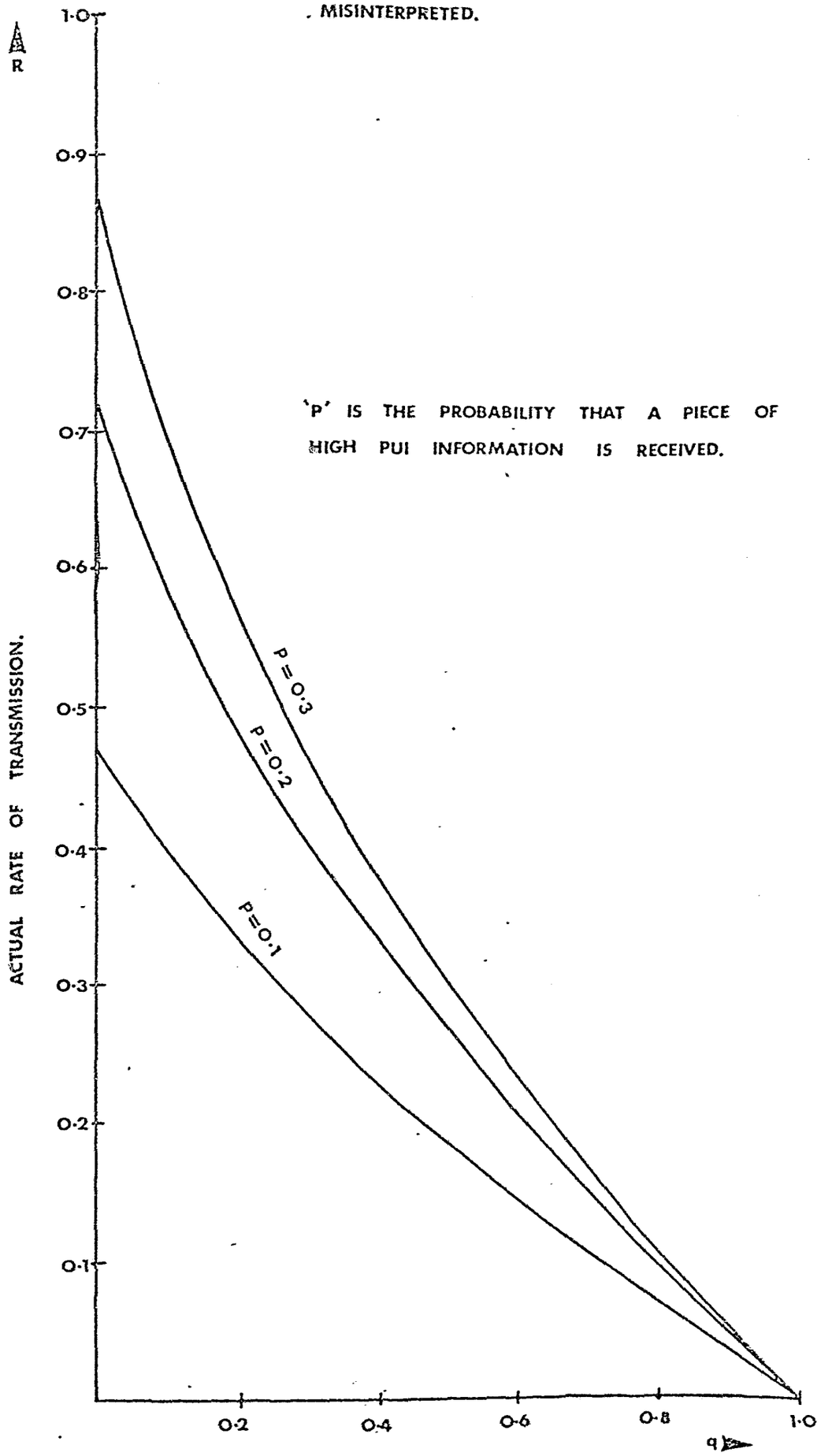


FIG. 7.

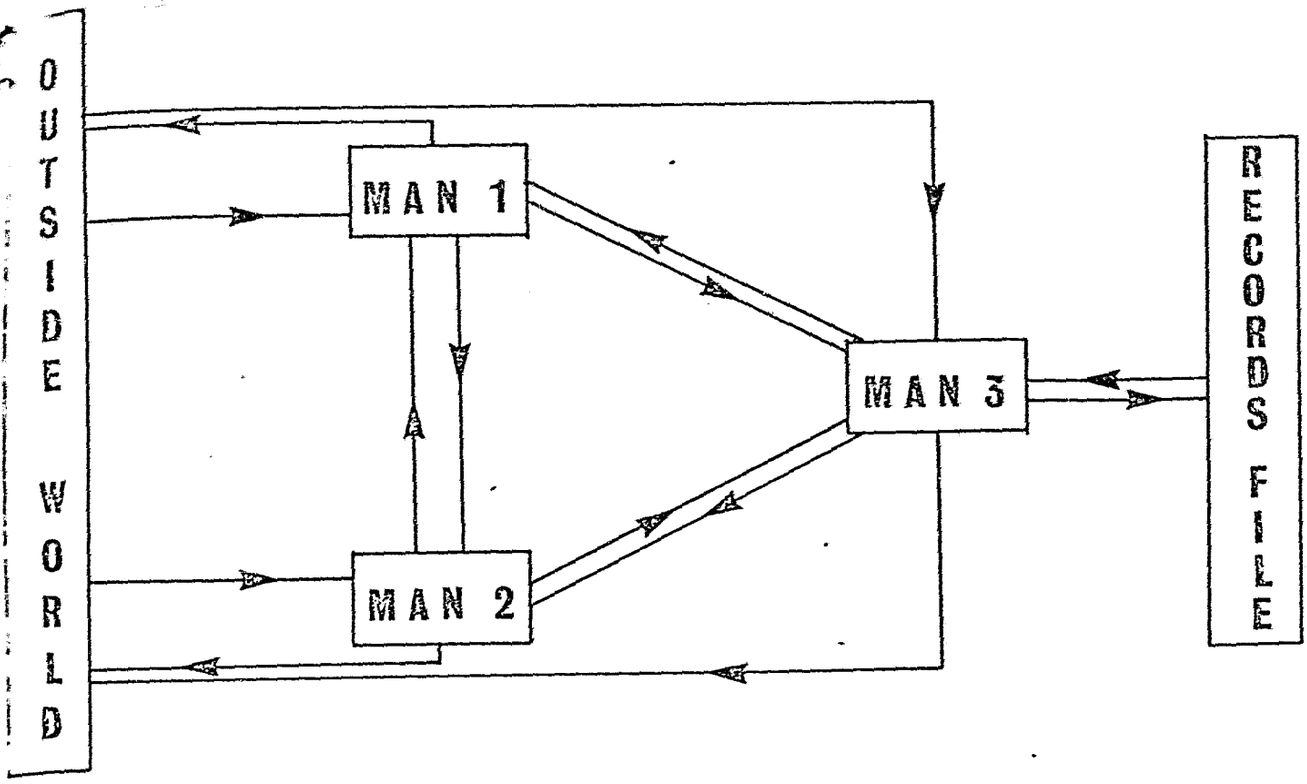


DIAGRAM SHOWING HOW INFORMATION FLOWS IN A TYPICAL POLICE STATION TODAY.

FIG. 8a

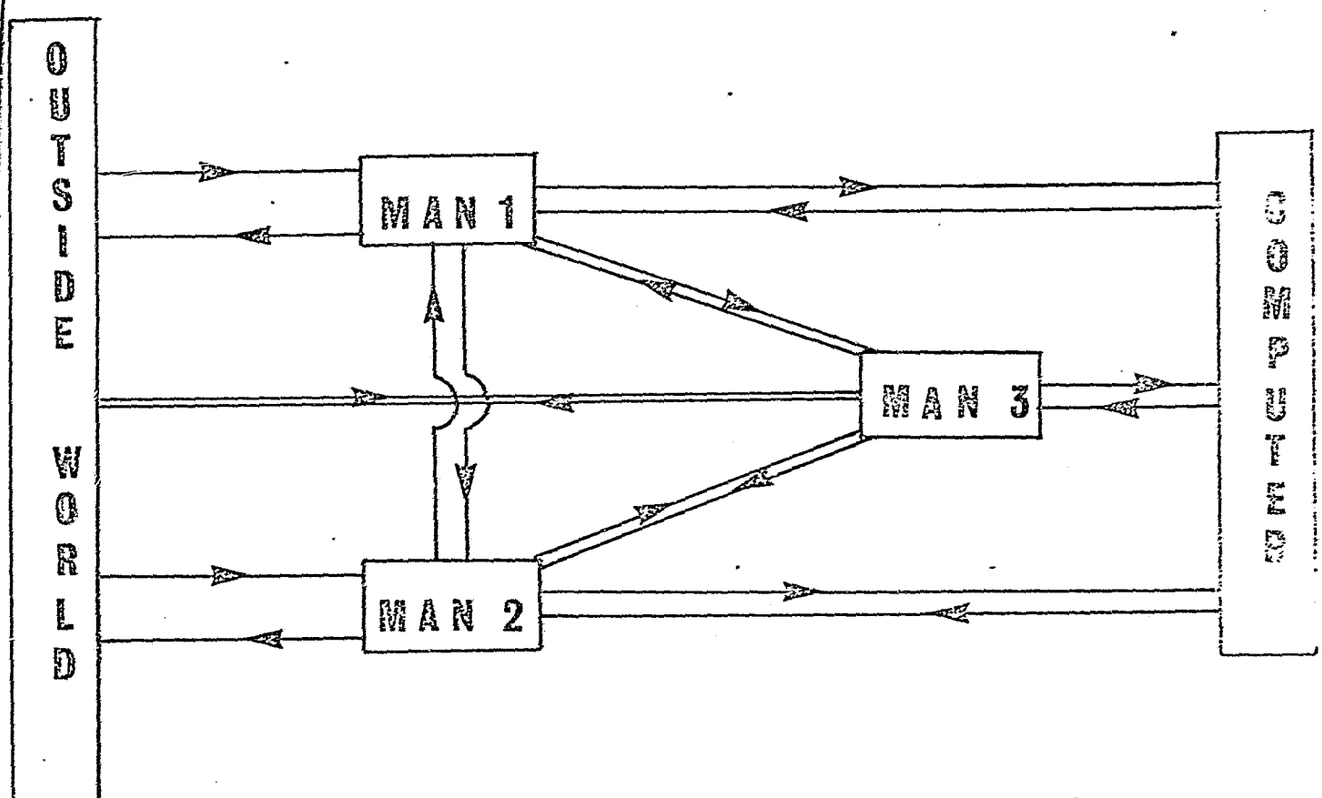


DIAGRAM SHOWING HOW INFORMATION COULD FLOW IN A SYSTEM WITH A COMPUTER.

FIG 8b

END

7 [illegible]