Information Evaluation in the Military Domain: Doctrines, Practices and Shortcomings

Philippe Capet, Adrien Revault d'Allonnes

Chapter 4

Information evaluation in the military domain: doctrines, practices and shortcomings

4.1. Introduction

Information evaluation, the fundamental theme of this book, is by no means a new idea, and is profitably used in a number of domains (physical, economical, biological, demographical and many more). Yet the establishment of a method, and the attempt to establish a certain rigor of this practice as part of an overall process, are, without a doubt, attributable to the domain of Defense. The origin of information, as well as its content, plays a crucial role both in a strategic and tactical context. Far from being reserved to espionage or data encoding, information evaluation is quite rightly considered to be an essential step in the preparation of a maneuver, the understanding of a situation in a theater of operations, or the making of a decision at a politico-strategic level.

In this chapter, we begin by presenting the doctrines**Error! Bookmark not defined.** in force regarding information evaluation (which does not necessarily mean that they reflect the real-world uses and practices of military intelligence personnel). Then, from a purely conceptual point of view, we hold up some of the shortcomings of these definitions and their associated underlying uses. These various theoretical difficulties could potentially cause very damaging practical consequences, in that they make any attempt at information evaluation impossible and, therefore, may have repercussions throughout the process of which information evaluation is a part. These open-ended issues, which are by no means exhaustive, will serve us in putting

---

Chapter written by Philippe CAPET and Adrien REVAULT D'ALLONNES.

forward possibilities for solutions, which will certainly need to be further developed, but which outline the work which remains to be done. The fictitious scenario presented in the introduction to the book is use**Error! Bookmark not defined.**d for three illustrations: the presentation of the existing concepts, some of the pitfalls relating to these, and the avenues opened up by the initial suggestions. The chapter refers directly to the previous one, focusing on points about information evaluation that were not expanded upon in Chapter 3, and also serves to establish useful perspectives for several of the chapters which follow.

4.2. Presentation of the existing situation

Information evaluation is essential to intelligence, helping guide highly important decisions by handling sensitive information. Hence, it is an integral part of the process of validation of information and, therefore, is subject to directives and guidelines.

Before describing the habits in force, it should be pointed out that, historically, such evaluation plays a part in the management of situations of open conflict – i.e. when surveying battlefields. However, the functions and requirements of intelligence have evolved; nowadays they are geared more towards peacekeeping, monitoring the stability of regimes or asymmetrical conflicts such as the struggle against terrorism. Therefore, the requirements and the materials have changed. First, because their perimeter is less restrictive than the identification of troops mobilized in a conflict zone, the analysis of clandestine groupuscules or political forces requires in-depth investigation. The appropriate reactions can only be determined once these investigations have been carried out, when knowledge has been constructed and established. Nowadays, the volume of sources has burgeoned, and the urgency to respond is less critical than in a traditional context. Decisions are made on the basis of more information, possibly gleaned from less stringently identified sources. Furthermore, the near impossibility of manually handling such volumes of data gives rise to a need for computer assistance in the processing of the information, as we can no longer solely rely on the expertise of a human operator.

Here, we present the existing doctrines**Error! Bookmark not defined.** and the underlying concepts, which are supposed to dictate the practices in force. The limitations of these practices will then be used to envisage the changes that are needed – particularly the clarifications and formalizations which are indispensable for semi-automated handling of information evaluation.

4.2.1. *Information evaluation in the intelligence cycle*

The activity of intelligence services – military, economic, public or private – is the series of operations whereby, following an initial request, information is gathered, assembled and then enriched and, finally, made available to the client [MIN 01]. When the response is given, it is possible that new questions will emerge and the process will have to start all over again. For this reason, we speak of the intelligence cycle.[1]

This cycle varies slightly according to cultures and customs, and therefore exhibits a variable number of stages. The version presented below, in Figure 4.1, includes the basic four stages found in all formulations of the model, which we briefly outline below, basing our explanations on the available doctrines**Error! Bookmark not defined.**.

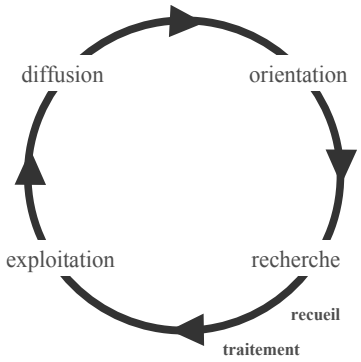

Figure 4.1. *The military intelligence cycle (simplified view)*

The initial phase – direction – is where the end user expresses his needs. In the historical context of military intelligence mentioned above (the surveillance of battle fields), the commander establishes zones of interest and the plan of collection – i.e. the list of services to be requested. The working plan thus constructed provides the direction for the search for information.

Second, the phase of searching for intelligence, carried out by human agents of the services in the plan of collection constructed during the previous phase, begins with the search for relevant sources. This collection of information procures the raw data upon which the following phases of construction of knowledge are based.

1 See Chapter 3 for further discussion of the cycle.

Exploitation – and the processing which immediately precedes it – are carried out to enrich the raw data thus extracted. It is mainly this phase which is of interest to us here. The five tasks which make it up, illustrated by Figure 4.2, are described here as they are in [TTA 01]. In France, the terminology has evolved slightly since 2001 – particularly with regard to information evaluation. We shall come back to this later. The first task is *grouping*, the task of classifying pieces of information of the same nature, i.e. relating to the same object. Then comes *information evaluation*, the evaluation of the quality of the information and of the sources consulted. This is followed by *analysis*, which aims to extract from the information those elements which are most significant for the response to the initial orientation. The next step is *fusion*, where the informational elements contained in different information items are integrated into one framework, in order to obtain enriched content. Finally, in the *interpretation* stage, we evaluate the reach – the set of consequences caused – of the information.

During the final phase – dissemination – the intelligence constructed in response to the initial question is given to the commander and to any agencies that it may concern. These addressees may then be led to redirect the search, triggering an iteration of the intelligence cycle.

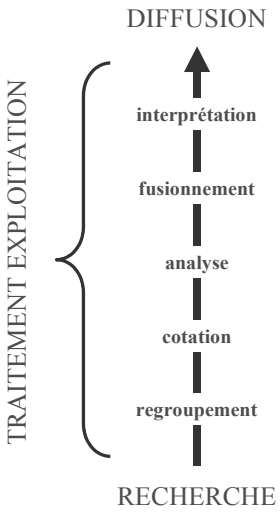

Figure 4.2. *Detailed representation of the exploitation phase in the intelligence cycle*

Thus placed, information evaluation can therefore be considered as a *task* during the *phase* of exploitation, fed by the *grouping* of information and feeding material to the *analysis*, i.e. a relevance filter for the next stage of the exploitation.

However, this placement of information evaluation within the process of intelligence production is sometimes called into question by other doctrines on the subject. Note, for instance, that the French inter-armed-forces doctrine [INS 03, p. 56] states that "information evaluation begins even during the first-level processing. It is validated or modified during exploitation […]". The inter-ally doctrine of the North Atlantic Treaty Organization (NATO – [NAT 03]), for its part, states that information *evaluation* is a stage in exploitation *stricto sensu* in the intelligence cycle, meaning that it takes place *after* processing. The French armed forces glossary of operational terminology gives information evaluation a similar place in the cycle: "a step in the exploitation phase of the intelligence cycle leading to an appreciation of raw intelligence in view of the reliability**Error! Bookmark not defined.** of the source and the credibility of the information" [DEF 04]. The question thus arises of the exact place of this activity in the intelligence cycle. As this question determines the object of information evaluation, i.e. what it relates to, it raises the problematic point of the granularity of the piece of information – either a piece of raw data or an enriched piece of intelligence – which is touched on later on in this chapter.

4.2.2. ***Reliability and credibility****Error! Bookmark not defined.* ***of information***

In the task of information evaluation which takes place after grouping, which in turn follows on from the search phase, the agent in charge of exploitation uses a two dimensional scale to evaluate the *source* of the information, on the first dimension, and its *content*, on the other. According to the NATO doctrine**Error! Bookmark not defined.** [NAT 03] for a given piece of information, the scores of the **Error! Bookmark not defined.***reliability****Error! Bookmark not defined.*** of the source and the *credibility****Error! Bookmark not defined.*** of the information are measured on the two six-graded scales shown in Tables 4.1 and 4.2 respectively. Note that this official classification is the same in France as that given by NATO, which explains the meaning of each grade in the right-hand column of the two tables. Up until quite recently, France's armed forces used terms which are slightly different to those used now: the *qualité* (quality) of the source of the information and the *valeur* (value) of the content. Certain definitions of the scores have also changed with a view to harmonization with NATO's glossary.

A	Completely reliable	Refers to a tried and trusted source which can be depended upon with confidence.
B	Usually reliable	Refers to a source which has been successful in the past but for which there is still some element of doubt in a particular case.
C	Fairly reliable	Refers to a source which has occasionally been used in the past and upon which some degree of confidence can be based.
D	Not usually reliable	Refers to a source which has been used in the past but has proved more often than not unreliable.
E	Unreliable	Refers to a source which has been used in the past and has proven unworthy of any confidence.
F	Reliability cannot be judged	Refers to a source which has not been used in the past.

Table 4.1. *Reliability of the information source: grade, label and description [NAT 03]*

1	Confirmed by other sources	If it can be stated with certainty that the reported information originates from another source than the already existing information on the same subject, then it is classified as "confirmed by other sources" and rated "1".
2	Probably true	If the independence of the source of any item of information cannot be guaranteed, but if, from the quantity and quality of previous reports its likelihood is nevertheless regarded as sufficiently established, then the information should be classified as "probably true" and given a rating of "2".
3	Possibly true	If, despite there being insufficient confirmation to establish any higher degree of likelihood, a freshly reported item of information does not conflict with the previously reported behaviour pattern of the target, the item may be classified as "possibly true" and given a rating of "3".
4	Doubtful	An item of information which tends to conflict with the previously reported or established behaviour pattern of an intelligence target should be classified as "doubtful" and given a rating of "4".
5	Improbable	An item of information which positively contradicts previously reported information or conflicts with the established behaviour pattern of an intelligence target in a marked degree should be classified as "improbable" and given a rating of "5".
6	Truth cannot be judged	Any freshly reported item of information which provides no basis for comparison with any known behaviour pattern of a target must be classified as "truth cannot be judged" and given a rating of "6". Such a rating should be given only when the accurate use of higher rating is impossible.

Table 4.2. *Credibility of content of information: grade, label and description [NAT 03]*

This double scale thus attaches to a piece of information a couple of values, referred to as a *rating***Error! Bookmark not defined.**, which are supposed to report the results of the evaluation of that information in order to be able to analyze it in the strictest sense. For instance, the information transmitted after direct observation of a tank by a combatant[2] can be rated B2, to express the fact that the soldier providing the information is certain of the fact, but may, in the eyes of the receiver of that information, be mistaken. The rating B2 means, however, that it is probable that a tank was in that particular place in the observed situation.

At first glance, in view of these definitions and their associated examples, information evaluation is firmly embedded in the intelligence cycle, and can officially and unambiguously be put into practice. However, is it all that certain? The next section of this chapter looks at the limitations of these directives.

2 Example taken from [TTA 01].

4.3. Illustrative scenario with multi-sourced information

The methods set out in the doctrines seem accurate: with the definitions and criteria formulated here, an officer charged with evaluating a piece of information in terms of its reliability**Error! Bookmark not defined.** and credibility**Error! Bookmark not defined.** has a circumspect and clear framework upon which to base his evaluation. Using the example of the fictitious scenario described in the introduction to this book, enriched with elements that are particular to this chapter, we are going to see the ambiguities and difficulties – both conceptual and practical – that still arise during the process of information evaluation, rendering it impracticable or devoid of meaning and applicability, if we adhere to the original doctrinal model. Let us return now to the case of the bomb attack perpetrated in the capital of Ektimostan**Error! Bookmark not defined.** on 31 May, and attempt to examine it through the eyes of, and through the lens of the knowledge held by, the agents of Usbek, a dissident now living in exile in a foreign country, and an unofficial supporter of the rebellion. Many different pieces of information, of diverse natures, are considered by these agents:

a. on 1 June, Captain Ixil, head of the Free Resistance of Dagbas (FRD), denies that his rebel movement is responsible for the attack;
b. on 31 May, in the wake of the attack, the minister Balldar, chief of the Ektimostanian police**Error! Bookmark not defined.**, *conditionally* incriminates the FRD;
c. the Ektimostanian head of State**Error! Bookmark not defined.**, Colonel al-Adel, states with certainty, on 2 June, that the FRD is behind the attack;
d. previously, a spy in the pay of Usbek who is part of the Ektimostanian police communicated to his interlocutors linked to Usbek**Error! Bookmark not defined.** that anti-governmental political groupuscules had been identified, were under heavy surveillance and were preparing to carry out violent actions against the regime, but are unconnected with the rebellion;
e. finally, on 3 June, a post by a very active and very widely read Ektimostanian blogger supports the idea that the FRD is behind the attack.

The sources of information are varied in nature: some of the reports are obtained by human intelligence *via* the spy in place (case *d*), others by intelligence of open-source origin such as the blogger (case *e*). The others are public declarations, so are also open-source, but with the peculiarity that they are relayed by press agencies or usual or authoritative Websites.

In reality, the information under examination refers to the same event with the five chosen sources: according to three of these sources, it is *true* that the attack was carried out by the FRD (this proposition is denoted as "ϕ" hereafter); according to

the other two, it is *false* that the FRD are the perpetrators of this attack – in other words, for these two sources, "not-ϕ" is true.

A number of elements from this scenario are used later on in this chapter to illustrate the limitations of the doctrines**Error! Bookmark not defined.** presented above.

4.4. From an inaccurate definition to an attractive but unusable concept

Now let us look at the way in which information evaluation is put into practice in order to gain a better understanding of its limitations. Significant criticisms can be leveled at the existing system for information evaluation. Indeed, before even thinking about whether or not the information evaluation accurately represents the realization of the fact that it qualifies, let us look at the way in which it can be apprehended – its intelligibility. Problems emerge with regard to each of the following four points: the reliability**Error! Bookmark not defined.** of the source, the credibility**Error! Bookmark not defined.** of the information, the combination of these two aspects to form a rating and, finally, the granularity of the objects being handled, ranging from raw information to enriched intelligence. Below, we shall examine each of these points in turn, and illustrate the limitations using the fictitious example presented in the introduction**Error! Bookmark not defined.**.

4.4.1. *Estimation of reliability*

4.4.1.1. *Reliability of the source: a question of point of view*

The use of the information evaluation scale is reliant upon the hypothesis that the first dimension, which only describes the source, is independent of the information. Viewed thus, the value of the reliability**Error! Bookmark not defined.** is supposed to be stable for all the information the source provides, independently of its content. However, according to the usage recommendations [DIS 01], it must reflect the trust that the analyst has in the source but can also indicate the conviction that the source has in the information that he offers, as well as his capacity to judge what he is providing. Indeed, all these elements seem judicious for the evaluation of the trust that can be put in the information, but they run counter to the linguistic labels in Table 4.1, which, for their part, describe only different levels of reliability. This confusion obscures this first dimension, leaving users to fall back on standard scores. Thus, as we saw earlier, a reliability score of B will be attributed to a field agent who is certain of what he has seen, or to a technological sensor. The top score, A, is not given so as to allow for possible failures on the part of the man or the machine.

It is therefore implied by the reliability scale that the source is rated regardless of his domain of expertise. The degree of trust that we are supposed to have in that source should normally be immutable. However, things almost never work that way: we give our trust[3] to a given person in a particular domain; less so in another, or perhaps more so, depending on the experience or knowledge that we have had of that source (I trust my plumber to repair a broken pipe, but not necessarily so much to remedy a power outage). The Ektimostanian blogger may be very knowledgeable about the economic environment, but no better informed than the man on the street with regard to the political issues in his country. The pro-Usbek spy may specialize in things financial or military in Ektimostan**Error! Bookmark not defined.** without knowing anything about the country's strictly interior policy.

Apart from the transmitter of the information, the receiver also has his own specific areas of expertise: how is an intelligence analyst specialized in the domain of economics to adequately judge the credibility**Error! Bookmark not defined.** of a piece of information in a domain which he knows nothing about, such as Ektimostanian politics and, directly related to reliability, grade a source whom he has never dealt with, or with whom he has dealt in circumstances which have misled him as regards to the level of trust that can be invested (e.g. not knowing that the same source had provided erroneous information on the subject on other occasions)?

The last impediment to the universality of the reliability of the source also relates to the receiver. Indeed, the evaluation of the source depends, as we have already stated, on shared history. This history is unique to the auditor. In addition, the auditor's subjectivity – or his allegiance – also influences the measurement. In fact, it is desirable that Usbek's intelligence services should not evaluate the reliability of Colonel al-Adel in the same way as his subordinate, Minister Balldar, would.

4.4.1.2. *Implementation*

The reliability of the various sources obviously depends on multiple factors: traditional objectives of the source as estimated by the intelligence agencies in the pay of Usbek, the antecedents, proclaimed or hidden interests relating to the stated information, etc. In a simplified approach[4] taking account of some of these factors, based on the definitions of the scale used to evaluate the sources, there is no real doubt about the reliability of the sources mentioned, depending on the adjustments which need to be made in the specific context.

In case *a*, the rebel Ixil sometimes denies actions which his movement has undeniably carried out and, conversely, vainly attributes to his own followers acts of

3 See Chapter 2 for an in-depth discussion of the notion of trust.

4 See Chapter 5 for further detail about methods for evaluating the reliability of sources.

violence for which others are responsible, or even which have never taken place. He is estimated to be "not usually reliable" and his reliability is evaluated at D.

In cases *b* and *c*, the Ektimostan leaders can be reliable, but have also been known to be propagandists. For example, it has been proven in the past that their declarations were true, exaggerated, played down or quite simply barefaced lies. Usbek's agents could therefore attribute them a score of C, based on the (highly debatable) principle that the ambivalence of the mode of declaration of the two personalities yields that average.

In case *d*, the active spy is considered to be a safe bet, because of his past history and his status with Usbek's intelligence agents. His reliability is therefore maximum, and is rated A.

Finally, in case *e*, the blogger has in the past proved his excellent knowledge of Ektimostanian environments and his blog seems to be objective, although it tends to lean in favor of the regime. Our agents therefore find him to be "usually reliable" – a reliability score of B.

Table 4.3 recaps the evaluation of the reliability of the various sources by the pro-Usbek agents, and recalls the position advanced by each one.

Information source	**Ixil**	**Balldar**	**al-Adel**	**Pro-Usbek spy**	**Ektimostanian blogger**
Reliability score	D	C	C	A	B
Proposition	not-ϕ	ϕ	ϕ	not-ϕ	ϕ

Table 4.3. *Reliability of sources, evaluated approximately on the basis of their past history and the knowledge that Usbek's agents have*

4.4.2. *Estimation of credibility*

Once the reliability of the sources has been evaluated, the credibility of the proposition remains in doubt. By examining various questions among possible others, we will now see that this credibility proves even more difficult to evaluate than reliability.

Similarly to that of reliability, the interpretation of credibility poses a problem. This dimension is supposed to represent the degree to which the information is

credible, as indicated by the labels of the final four predetermined levels on the scale (graded 2 to 5). However, because the maximum level is reserved for information "confirmed by other sources" – not to mention the problem of information granularity, to which we will come back – this is more an indicator of *confirmation* than of *credibility*. Section 4.4.2.1 looks particularly at this point.

4.4.2.1. *Acceptance and debatable role of corroboration*

Of the two contradictory events related by propositions φ and not-φ, which are we to believe? Do we believe the positive information transmitted by the three sources of middling reliability between B and C, or the negative information transmitted by the two sources whose reliability scores are very different (one is A, and the other is D)? How do we evaluate the credibility of that information, in view of these very different and heterogeneous reliability scores? As the information φ appears to be more widely corroborated than the contrary, it would be legitimate to give it the score of 1.

Indeed, level 1 in terms of credibility indicates that a piece of information is corroborated by several sources: when we receive information from one source, we note that its content is identical to that of a piece of information received previously from a different source. Its (maximum) rank indicates that therefore the information is as credible as it can be, and we can assume that the more sources agree on a piece of information, the more credible it will be judged to be. However, what can we actually conclude from a corroboration, without any other specification? We can easily see at least three difficulties in the definition and use of corroboration.

First, in our example, the spy seems to be a better choice to provide information to Usbek, for whom he works from within the Ektimostanian system, unlike the political personalities and blogger, who are all Ektimostanian. Yet the spy is the only one (besides a doubtful individual) who claims that φ is false. The statement that the attack was perpetrated by the FRD is corroborated by three individuals, whom one might imagine to be even more numerous, unlike its contrary which, if we ignore Ixil, who is too unreliable in the example, has no corroboration. Should we infer from this that φ is true?

It is not coincidental that the doctrine has come to include the reference to a *convention* of use of the score 1 when the source is marked A. Indeed, "by convention, the evaluation of A1 is reserved for exceptional cases when doubt is impossible" [INS 03]. Yet apart from these "exceptional cases" where corroboration no longer appears, *any* corroboration does not, in itself, indicate anything about the credibility of the information. For instance, a rumor may be shared and propagated through a crowd without its content necessarily being further validated. The voice of

the people (*vox populi*) is not necessarily the voice of an omniscient being *(vox Dei*); the common affirmation of a crowd is not necessarily credible, even if there is only one person who says the opposite.

Second, a piece of information and its opposite may both be corroborated. In our initial example, we have an almost equal proportion: three sources against two support the proposition ϕ, i.e. two contradicting corroborations – for ϕ on the one hand, and for not-φ on the other. The easiest way to resolve this ambiguity would be to impose a condition, so that in order to obtain a score of 1, the information must be confirmed by a large majority of sources providing information about that subject, and undermined by few or no sources, with the use of a threshold to be determined for the proportion of corroborations. As with the previous point, the rate of high reliability for such a corroboration, as opposed to a low average reliability for the contrary sources, needs to be taken into consideration. Yet even so, multiple possibilities would subsist, making it impossible to make a clear judgment: the question is more complex than it is with a vote, where the majority wins, because here, the "election" also needs to consider the reliability of the voters. Hence, a vote would need to be weighted by the reliability of the person casting it, using a formula which remains to be defined.

Finally, the last problem, the relationships between the information sources corroborating a piece of information ought not to be ignored: if two enemies are promulgating the same information, this means something different than if it were two allies, two people connected by the same interests. In our fictitious example, what advantage do we gain by noting that Balldar supports the same statement as his boss? Rather it is the opposite that would be surprising and worthy of interest. In reality it is more a question of *redundancy* than meaningful corroboration, and that redundancy should add nothing to the evaluation of the credibility of the information. Similarly, if we consider that the pro-Usbek agent shares neither Ixil's positions nor his objectives, the coincidence of their affirmation is informative *because of the fact* that they are supposed to be possible adversaries, and not in itself and independently of the pre-existing relations between sources. However, that coincidence (which resembles a corroboration) says nothing about the credibility of the information in itself; it informs the analyst, or at least piques his interest, about something entirely different. It is highly preferable for the sources corroborating a piece of information to be *independent* if its credibility is to increase; yet this is not what is said by military doctrine, where *any* corroboration seems to entail an increase in the credibility of the information.

4.4.2.2. *Problems in accessibility and unused top levels*

These difficulties in the apprehension of credibility**Error! Bookmark not defined.**, which may be at the root of the lack of objectivity in the usage and consistency between operators, are, without a doubt, partially resolved by the skill of the specialists. However, many of these operators struggle to avoid these problems. In cases where credibility can be evaluated, only the four levels qualifying doubt (grades 2 to 5) are used, and the integration of the confirmation (level 1) is left up to the services responsible for analysis, which perform matching prior to fusion of various pieces of information, once again posing the problem of the granularity.

In addition, from a purely semantic point of view, other problems arise with the evaluation of credibility**Error! Bookmark not defined.**. In our example (point *b* above), Balldar suggests *conditionally* that the FRD is to blame. In reality, how are we to score the credibility of the transmitted information? Balldar himself gives a sort of score to the statement that the FRD is to blame (it *could be involved*, he says: a score of around 3 on the scale). In view of this moderate presentation, we should not assign the same score to the ensemble of the information (source and content) as if Balldar had declared the same thing with categorical certainty, or if he had been highly doubtful but stopped short of rejecting that hypothesis. The source remains the same in all three cases; the event (*stricto sensu*) has not changed; but the overall score should change depending on the certainty or doubt expressed by the source.[5] In other words, semantic elements included in the information immediately alter the credibility that should be attached to it.

Furthermore, this problem cannot simply be rectified by incorporating a sort of "doubt coefficient" provided by the source itself, because linguistically speaking, many other nuances would then need to be given similar consideration. There is clearly a difference in meaning between "the Minister stated that a given event could have taken place" and "the Minister could have stated that a given event has taken place": the doubt expressed by the conditional is applied in turn by the source – i.e. the minister – and by the press agency reporting his comments. Yet in a strictly by-the-book approach to information evaluation, these two expressions should be taken to attach the same level of credibility to the same event, although the doubt is not expressed by the same source. Here, we touch on the tricky question of consecutive sources (and of the independence between source and content when evaluating each of them respectively): one reporting that another has suggested that a third party may have said that… etc. This question is specifically dealt with in Chapter 8.

---

5 See Chapter 7 for details about how allowances are made for this uncertainty expressed by the source, by a semi-automated method of evaluation.

4.4.3. *Combining dimensions – what is the comparability of the ratings?*

Information evaluation on a double scale, in addition to ensuring it properly represents the desired dimensions, is supposed to enhance the immediate readability of the rating. It seems, however, that operators are not all of that opinion: in addition to the difficulty in assigning a level of trust, the problem with readability of the resulting score also stems from a lack of comparability. Indeed, it is difficult to tell which piece of information is more credible if one is scored B3 and the other C2 – two levels which should, in theory, be comparable. The specialists with whom we have been able to speak on this subject maintain that only the known values, i.e. the "default values" often cited here, are expressive and therefore interpretable. A piece of information gathered by a drone, which is therefore considered to be theoretically certain, is evaluated as B2, establishing a form of benchmark, which is unfortunately impossible to compare against other, less objectively envisaged situations.

Another weak point of such an expression of the evaluation, as a combination of scores, stems from the interpretation of the credibility**Error! Bookmark not defined.**. Indeed, construed as an indicator of confirmation, the scale it uses runs from invalidation to confirmation. Thus, it is a signed (positive and negative) scale, constructed around a neutral value. Indeed, two pieces of information may more or less confirm or undermine one another, or they may be completely unrelated. Reliability**Error! Bookmark not defined.**, on the other hand, describes a progression from low or non-existent activation to an absolute maximum. The combination of values on these two differently constructed scales increases the margin for interpretation by the users and, along with it, the risk of inconsistency between their perceptions.

With Tables 4.1 and 4.2, the intermediary ratings (B-D and 2-5) use adverbs and adjectives which are supposed to characterize the different levels: *usually*, *fairly* and *not usually* for reliability**Error! Bookmark not defined.** and *probably*, *possibly*, *doubtful* and *improbable* for credibility**Error! Bookmark not defined.**. With regard to reliability, these terms correspond to an approximate average of the credit attributed to a source after the analyst's findings; with regard to credibility, to a level of belief of the analyst on the basis of his expertise and his experience. While the order of these words is more or less undisputable because there is little risk of its varying from one analyst to another, the boundary between one class and another may be extremely variable with different experts. Quite apart from the subjectivity of the analyst, the choice of a level, with corresponding opinions, depends on the sensitivity of each analyst to the qualifiers used.

In our example, one analyst may well estimate that the Ektimostanian blogger is "usually reliable", whereas another would say that he is "not usually reliable",

depending on the particular analyst's past experiences of reading the posts on his blog, and his particular knowledge, but also on the subjective expectations that each analyst has: even supposing that the two analysts have read the same things on the blog and have the same knowledge of the domain, one may employ the adverb "usually" where the other would use "rarely" (or "not usually"), and the rating is therefore merely a question of point of view.

As there is no other criterion that is more objective than this use of subjective terminology, information evaluation is "fuzzy"; the same is true of credibility. However, the two scales for information evaluation have only six degrees each. These *discrete* grades are based on fuzzy adverbial estimations which in fact more closely reflect a certain *continuity*. The multi-valued logic with 6 values of truth implicitly present in these scales expresses an assumed position which fuzzy logic [BOU 93], or "Computing with words" [ZAD 02], seems better able to express if we wish to preserve these adjectives and adverbs and the gradation of the scoring system attached to them.

4.4.4. *Raw data, enriched intelligence – can information evaluation qualify everything?*

As previously stated, depending on the particular doctrine in question, information evaluation takes place at different stages in the intelligence cycle. The French manual for joint forces [INS 03], for instance, specifies that it is applied to information[6] rather than to intelligence, although the producer of intelligence does pair it with an evaluation. A distinction indeed needs to be drawn between information – facts *reported* to the agencies by a source – and intelligence – the agglomeration of knowledge from relatively diverse sources, *produced* by the agencies. In addition, during the enrichment of the data and the production of intelligence, it is clearly stated that these two types of data are mixed, so a piece of intelligence may be built on raw data and other pieces of intelligence. The choice not to measure trust in the same way for these objects of different granularities is, doubtless – at least in part – a consequence of the choice of scoring system. Indeed, the notion of source reliability poses a problem for a piece of intelligence, because it is the product of the fusion of different snippets of information from different sources. However, the evaluation of the sources is still relevant in order establish trust in a piece of intelligence.

In addition, as these evaluations are not, as we have seen, comparable, the construction of a score for a piece of intelligence by aggregation of the scores for the

6 Here, for simplicity's sake, we count the raw data to be information.

different pieces of information is not a clearly formalized operation. In practice, the analyst responsible for that aggregation demonstrates his skill by writing up accompanying remarks. These operations are further complicated when creating intelligence that combines pieces of information and other intelligence, which itself has been created by the fusion of data. The analyst then forms an opinion about the trustworthiness on the basis of different types of indicators, which are therefore incomparable. The skill of the analysts is, of course, unquestionable. However, this simply makes the formalization of information evaluation, either to facilitate its automation or to improve its readability, more confusing. Suppose that, rather than raw data, the report of the pro-Usbek agent forms part of an enriched piece of intelligence; where and how does his reliability fit in to the overall evaluation of that intelligence?

If we look at the problem of granularity, we see the inconsistency of the top level of credibility. Indeed, the maximum degree of credibility of a nugget of information depends on confirmations – or invalidations – by other sources. As information evaluation takes place before grouping and fusion, taking into account other pieces of information runs counter to the definitions of a given atomic element and its evaluation. Note also that this divergence introduces the combination of sources, which is difficult, as we have just pointed out. While this observation seems to provide an argument for abandoning corroboration when evaluating the quality of information, the search for confirmation to relieve a doubt is so widespread and so often appropriate that we are led to stick with this approach.

Section 4.5 briefly runs through some of the publications on the subject, before the rest of the book goes on to detail some of these proposals and offers new ones.

4.5. A few suggested refinements to information evaluation techniques

Military research in general, and research into issues relating to intelligence in particular, are experiencing rapid growth. While information evaluation is not the most widely studied domain in this context, in this section we introduce a number of articles which deal specifically with the issue as defined above. Our introduction to each of the articles on the subject contains some of our own remarks about the existing model of information evaluation. The proposed methods are usually anchored in the maintaining of consistency in a system or the combination of degrees of uncertainty, which are classic tasks of uncertain information fusion.

Thus, Cholvy and Nimier [CHO 03] focus on maintaining consistency in a knowledge base, and look at various operators that can be used for fusion, weighting the distances between possible worlds according to the reliability of the sources.

Cholvy [CHO 04] reuses the initial information evaluation grid to qualify the elements inserted into a database storing the history of the pieces of information and their sources. The responses to requests to the database are chosen by a vote between conflicting elements. This history contributes to the construction of a piece of information enriched by the fusion of the stored elements.

In the same vein, most of the articles discussed here consider the problem of information evaluation to be a fusion problem. Nimier [NIM 04], for instance, proposes a description of it using combination operators in three formalisms: the first probabilistic, the second possibilistic and the last based on Dempster–Shafer theory.

In greater detail, Besombes and Cholvy [BES 09] propose a complete architecture for an information fusion system. In order to perform fusion, the pieces of information are mutually correlated, the correlation measurement covering the whole of the scale, from invalidation to confirmation. The authors then introduce a calculation to update the evaluation on the basis of this correlation. Although they do not detail the fusion process, they examine different methods for calculating the correlation.

These articles thus envisage information evaluation as an estimation, in a more or less closed world, of the contradiction between known facts. We have seen that credibility can be seen as a confirmation indicator, essential in the evaluation of trust. However, we have also seen that to consider only that factor is not at all satisfactory, and is among the problems with the existing approach to information evaluation. In addition, these models, however rich they may be, assimilate the evaluation of a piece of information to its consistency with existing knowledge about the world, leading to disagreement as to the mode of production or summarizing it as the relative importance of the reliability of its source.

Another point of view is proposed by Baerecke [BAE 10], who, in addition to the dimensions usually taken into account, considers the confidence that the source attaches to his information, the recentness of the information and the amicable or hostile relations between sources. Uncertainty is given as distributions of possibility moderated by the reliability of the sources. Elements are then fused according to the friendship partition, that is two sources in conflict with one another but providing the same information corroborate one another more than two friendly sources doing the same.[7] This formulation reuses certain aspects of information quality [BER 99] to determine the factors needing to be taken into account in its evaluation. The

7 See Chapter 7 for an example of how to take account of these relations between sources in the information evaluation process.

authors add notions of subjectivity and propose an interesting model of the repercussions of their dissemination.

In [REV 11], we put forward a different view of information evaluation, whereby it is understood less as an evaluation of the reality of the fact described than as an estimation of the faith that can be invested in the information describing that fact. We will come back to these works in Chapter 9.

4.6. Conclusion and future prospects

The conclusion drawn from the above discussion should not be too pessimistic: the inclusion of information evaluation in the phase of exploitation of the information is crucial, and to deprive it of its role would be calamitous for any labor of intelligence, the information would then only need to be relativized, in the postmodernist sense where "everything is of equal value". However, we still need to know how to cater for the need for information evaluation in the restricted context of military intelligence, using proposals for reform of the two original dimensions, balancing the specific levels of each of them and a possibility of usage, even if it requires a prioritized enrichment of the doctrines of other dimensions of information evaluation.

Yet there are crucial questions which remain. Thus, even if these dimensions of evaluation are refined, the relationship between the source and the content needs to be far more clearly defined. In many scenarios, though, there is probably a certain degree of dependency between them, whether we are working in an overall framework or one relative to the information received. In addition, throughout the process, the human agents will inevitably remain subjective. This crucial issue, and that of the scales used for evaluation, merit in-depth theoretical studies using a variety of possible models (fuzzy logic, multi-valued logic, etc.), necessarily followed by tests using experts in the domain to refine the classification scales thus constructed. By doing this, the practice of information evaluation would be greatly facilitated, whilst the role of information evaluation would be enhanced throughout the phase of processing and exploitation of the information.

Indeed, if we look again Figure 4.2 representing that phase, and consider the decrease in the limitations to the methods and practices laid out in this chapter, not only would information evaluation continue to play the key role in the information exploitation stage to valorize it for military intelligence, but it could acquire an even greater role, throughout the whole length of this phase: at each step, from grouping, through analysis, fusion and interpretation to the beginning of the dissemination phase, it would likely be performed again each time the process advances to the next

phase, whilst helping the enactment of each one. Acquiring an additional status, at all times during the processing and exploitation of the information, information evaluation would thereby gain the respectability which it doubtless deserves.

4.7. Bibliography

[BAE 10] BAERECKE T. *et al*., "Un modèle de cotation pour la veille informationnelle en source ouverte", *Colloque international veille stratégique, scientifique et technologique*, *VSST'2010*, 2010.

[BER 99] BERTI L., *Quality and Recommendation of Multi-source Data for Assisting Technological Intelligence Applications*, Lecture Notes in Computer Science, Springer, Berlin, 1999.

[BES 09] BESOMBES J., CHOLVY L., "Information Evaluation in Fusion using Information Correlation", *International Conference on Information Fusion*, p. 264-269, 2009.

[BOU 93] BOUCHON-MEUNIER B., *La logique floue*, Presses Universitaires de France, Paris, 1993.

[CAP 12] CAPET P., DELAVALLADE T., "La cotation de l'information : approches conceptuelles et méthodologiques pour un usage stratégique", *La qualité et la gouvernance des données au service de la performance des entreprise*, Hermès, Paris, 2012.

[CHO 03] CHOLVY L., NIMIER. V., "Information Evaluation: discussion about STANAG 2022 recommendations", *NATO-IST Symposium on military data and information fusion*, Prague, Czech Republic, 2003.

[CHO 04] CHOLVY L., "Information Evaluation in fusion: a case study", *International Conference on Processing and Management of Uncertainty in Knowledge-based Systems*, *IPMU 2004*, Perugia, Italy, 2004.

[DEF 04] MINISTÈRE DE LA DÉFENSE, Glossaire interarmées de terminologie opérationnelle, Paris, 2004.

[DIS 01] DEFENCE INTELLIGENCE AND SECURITY SCHOOL & DEFENCE INTELLIGENCE AND SECURITY CENTRE, Intelligence Wing Student Précis, 2001.

[INS 03] INSTRUCTION INTERARMÉES SUR LE RENSEIGNEMENT D'INTÉRÊT MILITAIRE, TITRE I, Doctrine interarmées du renseignement, PIA 02-200, Paris, 2003.

[MIN 01] MINISTÈRE DE LA DÉFENSE. Commandement de la formation de l'armée de Terre (COFAT), Manuel du cadre de contact (TTA150), 2001.

[NAT 03] NORTH ATLANTIC TREATY ORGANIZATION, Standardization Agreement, Intelligence Report, STANAG n° 2 511, 2003.

[NIM 04] NIMIER V., "Information Evaluation: a Formalisation of Operational Recommendations", p. 1 166-1 171, *Seventh International Conference on Information Fusion*, Stockholm, Sweden, 2004.

[REV 11] REVAULT D'ALLONNES A., Evaluation sémantique d'informations symboliques : la cotation, Doctoral thesis, Université Pierre et Marie Curie, Paris, 2011.

[TTA 01] TRAITÉ TOUTES ARMES, *Renseignement*, n° 150, titre VI, Paris, 2001.

[ZAD 02] ZADEH L.A., "From computing with numbers to computing with words. From manipulation of measurements to manipulation of perceptions", *International Journal of Applied Mathematics and Computer Science*, vol. 12, n° 3, p 307-324, 2002.