***Weighing the Evidence: The BCISS Iraq HUMINT Analytic Matrix Exercise***

Professor Philip H.J. Davies[1]

Director, Brunel Centre for Intelligence and Security Studies (BCISS)

Brunel University

Uxbridge, Middlesex

UK, UB8 2DD

Email: philip.davies@brunel.ac.uk

Tel: +44 (0)1895 266 827

Dr. Kristian Gustafson

Deputy Director, BCISS

Brunel University

Uxbridge, Middlesex

UK, UB8 2DD

Email: kristian.gustafson@brunel.ac.uk

Tel: +44 (0)1895 265436

*…there is always a well-known solution to every human problem — neat, plausible, and wrong*

H.L. Mencken[2]

**Introduction: Education, Training and Understanding**

In the fourteen years since the Brunel Centre for Intelligence and Security Studies (BCISS) was established at Brunel University analytic methodologies and training have become an increasingly central feature of the Centre's activities. From its outset in 2004, the Centre's heavily subscribed MA in Intelligence and Security Studies featured analytic practice as the centrepiece of its programme through a model Joint Intelligence Organisation practical entitled the Brunel Analytical Simulation Exercise (BASE).[3] As the publicly available information and the state of the art in analytic professional practice progressed it became necessary to continually upgrade and expand the analytic offering. This was particularly true with regards to incorporating practical, hands-on exercises in the taught curriculum, ever more so as BCISS became more active providing continuing professional development (CPD) training for organizations like the European Union's Intelligence Centre (EU INTCEN), the British Army, and certain allied European agencies.[4] This has involved a combination of pre-existing exercises developed elsewhere and an assortment of home-grown practicals tailored to the particular pedagogical goals and philosophy of BCISS' teaching team.

One of our former colleagues at the Centre (now at The James Madison University), Stephen Marrin, has drawn a scrupulous distinction between intelligence 'training' and 'education'. The former 'is usually associated with government programs intended to provide the intelligence analyst with specific instructions to implement job-related tasks' and the latter refers to 'academic courses or programs geared to provide the student with knowledge along with conceptual and theoretical frameworks useful for understanding and exploiting that knowledge'. From a practitioner's point of view, education yields longer term performance benefits with the value of training being short term.[5] To which one might also add that education is also the foundation required for academic scholarship on intelligence institutions, processes and events. This distinction is worth noting because on the Brunel programmes each practical is adopted or designed to do double duty as both training and education. Where specific analytic tradecraft skills and methods are taken up the learning objective as much to provide a deeper intellectual *understanding* of how intelligence works and how and why specific intelligence mechanisms or events have played out one way or another.

One practical exercise that we have developed at BCISS deliver examines the human intelligence (HUMINT) on Iraqi non-conventional weapons programmes available to the UK government through

the British Secret Intelligence Service (SIS) in the eighteen months or so prior to the 2003 invasion of that Iraq. The Iraq HUMINT Analytic Matrix Exercise is ordinarily expected to run between 90 minutes and two hours, and is delivered on both the residential and distance learning versions of our intelligence MA as well as the various CPD courses. It draws on the comparatively detailed information about SIS's stable of six 'main' sources in Iraq was provided by Lord Butler in his 2004 *Review of Intelligence on Weapons of Mass Destruction*.[6] In the exercise described here, that information is interrogated through a variety of analytic matrix methods combined with the application of certain, standardised source grading standards and conventions. This task has a broad range of learning goals. In the first place, it is intended to get students thinking analytic matrix methods and manipulation through from first principles before introducing them to the Palo Alto Research Centre's Analysis of Competing Hypotheses (PARC ACH) software tool. In part, this is intended to prepare them for intensive use of PARC ACH during the latter stages of BASE. In other part, it is also used to make students think through the challenges of weighting and integrating multiple sources of information with highly variable levels of reliability and probable truthfulness in making cumulative analytical judgements. It also has a concrete historiographic goal of making students realise just how difficult the analytical task confronting allied intelligence communities in the summer and autumn of 2002, and how and why it may have been so difficult to 'call it right' on Iraq's weapons of mass destruction (WMD). It also has an additional historical function of providing a point of entry into the origins and evolution of analytic matrices in particular, and structured analytic techniques (SATs) in general.

**The Evolution and Application of Matrix Methods**

For teaching on intelligence studies, two pedagogic game-changers that emerged from the Iraq post-mortems were the attention paid to analytic methods for testing hypotheses by the so-called 'WMD Commission' in the United States[7], and PARC releasing a downloadable version of its ACH tool 'to the general public at no cost when used for non-commercial or educational purposes'.[8] Previously, dedicated commercial software tools like i2[9] or Palantir[10] were (and remain) prohibitively expensive, while their principal open-source alternative Maltego[11] (originally designed to run via Linux) had some significant practical limitations. In any event, while these tools were excellent for addressing things like link or network analysis and the practicalities of intelligence databasing they did not really address the kinds of issues the really form the centrepiece of academic teaching on strategic intelligence. Driven largely by the literature of intelligence failure, intelligence scholarship tends to be more concerned with issues of the reliability (or not) of sources in raw intelligence collection, the many analytic pitfalls that dog finished intelligence production, and potential administrative and

procedural breakdowns in the machinery of the intelligence community.[12] Entirely apart from cost advantages, PARC ACH alone offered a software tool designed to address the first two problems explicitly, and in its emphasis on use by analytic *teams*, the latter at least implicitly. PARC ACH also has the advantage of being available both as a Windows application and in two platform-independent Java variations that do not require installation.

One of the present authors (Davies) had already subjected the information on the UK's HUMINT sources in the Butler report to detailed scrutiny in the context of the long-term organizational issues affecting the Secret Intelligence Service.[13] Thus, when PARC ACH became available in 2006 followed in 2009 the CIA's *Analytical Tradecraft Primer* which gave additional detailed guidance to ACH methodology[14] the teaching team at the Brunel Centre for Intelligence and Security Studies found itself in a position to combine the American materials ACH and wider analytic tradecraft with the details published by Lord Butler in an exercise that uses the general principles of analytic matrices as a means to both examine source validation and evaluation and also give students hands-on experience using software tools for intelligence analysis.

In fact, the principles and much of the methodology of analytic matrix manipulation and ACH is far from new. Indeed, the credit for originally pioneering a recognisable precursor to ACH almost certainly goes to the interwar pre-war French *Deuxieme Bureau*. Unfortunately, the only known account of the Bureau's method is an one provided of a deeply sceptical British officer, Kenneth Strong (soon to be Dwight Eisenhower's head of intelligence and future founding head of UK Defence Intelligence[15]). Visiting the Bureau on liaison on a number of occasions between 1939 and 1940, Strong was troubled by a number of aspects of the French intelligence effort against the Germans, none more so than 'that certain Cartesian scholasticism with which the French approach all intellectual problems'.[16] In fairness, the Bureau's methods were probably more explicitly rooted in the deep and lasting influence of Baron Antoine-Henri Jomini on French military thought rather than any vague, ambient cultural fixation.

Jomini has a real (if inadequately examined) significance to intelligence studies because he consistently displayed greater confidence in the value of intelligence and its methodical analysis than his Napoleonic contemporary and sometime intellectual rival Carl von Clausewitz[17]. According to Jomini:

> 1. A general should neglect no means of gaining information of the enemy's movements, and, for this purpose, should make use of reconnaissance, spies, bodies of light troops commanded by capable officers, signals, and questioning deserters and prisoners.

2. By multiplying the means of obtaining information; for, no matter how imperfect and contradictory they may be, the truth may often be sifted from them.

3. Perfect reliance should be placed on none of these means.

4. As it is impossible to obtain exact information by the methods mentioned, a general should never move without arranging several courses of action for himself, based upon probable hypotheses that the relative situation of the armies enables him to make, and never losing sight of the principles of the art. [18]

Jomini concludes that there is indeed a method to fill in, or at least bridge, those inevitable information gaps. 'A skillful general', he argues, 'may supply the defects of the other methods by making reasonable and well-founded hypotheses.'[19]

Strong was profoundly suspicious of substituting any formal processes for human judgement and insight, whether via the strictures of uncompromising logic or automation through data processing technology.[20] Nonetheless his account is worth quoting at length, according to which:

> Applied to Intelligence [sic] estimating, … [French Cartesian] concepts gave rise a theory of *hypothèses.* First, the situation on the opponent's side was studied in detail, and the known facts established as accurately as was possible. A list was made of the possible courses of action of which he appeared to be even remotely capable. These were known as *hypothèses.* Further examination was intended to reduce them to a sensible minimum, but no one course was rejected as impracticable unless definite information was available to suggest that the opponent had rejected it. Those remaining were yet again examined, this time in greater detail, with the ultimate object of arriving at a single course of action which it could be reliably estimated the other side would adopt. Unfortunately such a neat solution was seldom possible, and often French intelligence officers appeared reluctant to abandon even the most glaring improbabilities except on strictly logical or factual grounds.[21]

It would be very valuable to see a French account of this period and its methods, but to the authors' knowledge none has yet been forthcoming in either French or English. In any event, apart from Strong's recollections the French precedent appears to have been lost to the collective memory of the intelligence profession, in no doubt largely due to the disruptions and dislocations caused by the subsequent fall of France.

In warning intelligence, testing indicators of a specific critical incident – typically but not confined to a military attack – against alternative explanations is in many respects a form of matrix methodology

closely akin to ACH. In what Richards Heuer and Randy Pherson and his collaborators have termed the 'indicator validator' technique, one examines indicators of imminent attack might also appear as consequences of an escalated or more aggressive deterrent posture or even military mobilisation in support of a public emergency such as a natural disaster. In a civil contingencies emergency close to or just inside a contested border one might well see mobilisation of air reconnaissance to ascertain the scale of the emergency and locate survivors and routes of access to the affected area; military medical supplies to aid survivors; and substantial troops, military engineering and logistical capabilities to do the literal and figurative heavy lifting of evacuation and limiting the impact of the disaster. In Pherson's formulation of this task, indicators unique to a single scenario or likely only in a limited number of alternative possibilities would be considered strong or 'highly discriminating' indicators, ones that would occur in a large number of situations comparatively weak or 'nondiscriminating' indicators.[22]

This in many respects such indicator validation is already talking about classic components of ACH methodology such as testing alternative interpretations (each scenario being, essentially, a hypothesis) and judging the diagnosticity of specific items of evidence, that is, nominated indicators or indications actually received.[23] Indeed, the BCISS team has used a classroom 'audience participation' practical on indicators and warning (I&W) validation as a prelude to key analytic tradecraft concepts such as Douglas MacEachin's notion of 'drivers and linchpins'[24] and hypothesis matrices for the better part of a decade.

Within counter-intelligence a range of matrix methods have also evolved over the last couple of decades, particularly association and activity matrices.[25] Of the FBI's investigation into Aldrich Ames it has been reported that 'investigators had an idea of what was lost, then drew up a matrix of who had access to the information'.[26] Robert David Booth has also provided a detailed account of how a similar methodology was used to expose Kendall Myers as Cuban-controlled agent within the Department of State[27]. But these mainly sorting or collating tools designed to help analysts visualise and exploit intelligence reporting rather than being intended to articulate and test hypotheses as such.

Indeed, the counterintelligence use of matrix manipulation is a very good illustration of the fact that analytic matrices can be used in two very different fashions for rather different purposes. *Hypothesis* matrices are concerned primarily with falsification since experimental falsifiability is the essential scientific role of a hypothesis. Here the principal concern is with identifying and collating *disconfirming* evidence against possible explanations, interpretations or predictions. However, in a less Popperian fashion, matrices are also valuable *collating* tools, logging the full gamut of the

available reports as well as correlating them with possible implications and explanations. In this sense, the use of matrices in intelligence assessment is somewhat akin to double entry accounting in finance, assessing assets debited or credited against possible accounts where the items are evidence are the assets and the hypotheses or scenarios are the accounts. It also aids analysts in visualising reporting information in identifying correlations and more varied forms of pattern recognition. Advocates of ACH are fond of pointing to the limited buffer capacity of human short term memory[28], so using some sort of matrix is valuable for ensuring all of the available reporting is placed on the metaphorical table in the assessment process

**Pre-War Intelligence on Iraq as a Teaching Source**

In the UK, years following the invasion of Iraq might almost have been described as a self-flagellating and sometimes frenzied 'decade of intelligence'. In the House of Commons, the Foreign Affairs Select Committee got into a turf war with the joint Commons-Lords Intelligence and Security Committee (ISC) while conducting an inquiry that was more witch-hunt than post-mortem[29] while the ISC's own report on Iraq was notable more than a whiff of partisanship on the part of its Labour Chair.[30] The apparent suicide of a principal witness in the FAC investigation prompted a judicial inquiry[31] followed by a formally independent review by a committee of Privy Councillors led by Lord Robin Butler of Brockwell.[32] The Butler report was written in such nuanced 'Whitehall Mandarin'[33] that public dissatisfaction would not be mollified for a dozen more years until Sir John Chilcot, an alumnus of the Butler inquiry, led production of the 12 volume report of The Iraq Inquiry (recommended retail price £767.00/US$935.00).[34] Hard though it may be to believe, the Chilcot report actually added little additional detail on intelligence sources to that provided by Lord Butler, typically drawing directly on and referencing the Butler report on the question of raw intelligence.[35] Consequently the 2004 Butler report remains the principle source of information on SIS HUMINT in pre-war Iraq.

The advantage of Lord Butler's account is that it involves a manageable number of sources, six in total, with some relatively detailed validation information (quoted or paraphrased by Lord Butler from SIS source reports). The analytical problem – the hypotheses to be tested – is also of manageable proportions, and can be introduced as being essentially 'Saddam has chemical and biological weapons (CBW)' and 'Saddam does not have CBW' at the outset with additional hypothesis of greater nuance (such as a 'break out capability' of rapid WMD development in the event international sanctions were lifted[36]) being introduced as the exercise proceeds. Butler also offers enough information about the exploitation of these sources to give a reasonable idea of where their reporting pointed in terms of supporting either of the two possibilities.

That being said, trying to piece together a coherent picture of SIS' sources in Iraq is actually quite difficult because various parts of the HUMINT narrative in the Butler review are imperfectly consistent with other parts.[37] This problem is intensified by a tendency to switch between discussing specific reports and individual informants in different parts of the report without making it clear which source is being referred to when. Instead one has to rely on internal evidence and recurrent forms of words describing the sources. Thus while Lord Butler provides a concise account of the 'main' HUMINT sources it is apparent that there were additional minor or secondary sources providing limited reporting and typically through reporting chains of two or more cut-outs. Unhelpfully, there are also two *different* summaries of the SIS stable of informants on different pages[38] which sometimes duplicate one another verbatim and sometimes given completely different details organised differently on the same agents.

To make matters still harder to decipher, the cut-outs doubled as intelligence sources in their own rights reporting in their own specific spheres of competence. This, however, as we shall see below, this actually proves a very useful pedagogical item because this means that what appears to be a source on Iraqi WMD is actually acting only as a cut-out for a number of other sources.

For the purpose of this exercise, attention is focused on the six 'main' sources which can be summarised[39] roughly as:

- *Source 1* who 'reported accurately and authoritatively on some issues' but 'on production of stocks of chemical and biological agents, he could only report what he learned from others in his circle of high-level contacts in Baghdad.'[40] This appears to have included the claim that it was '"common knowledge" within his circle that chemical agent production was taking place.'[41]

- *Source 2* considered 'an established and reliable source' whose reporting 'on other subjects had previously been corroborated'[42]. However, this second source began to pass information received from a number of contacts who acted as a sub-agents or 'subsources', including, most notably:

- *Source 3* whose information 'properly included a caution about the subsource's links to opposition groups and the possibility that his reports would be affected by that'[43] and was subsequently withdrawn as unreliable.[44] It subsequently appears that this agent was the source of the notorious '45 minute claim'.[45] A number of items are credited to subsource reporting, most or all of which appears to have passed through Source 2. These include an account by a

source of being 'part of a project to produce the nerve agent VX in the period to 1998'; the concealment of 'up to 20' Al Saddam ballistic missiles[46]

- *Source 4* and *Source 5* were consistently grouped and discussed together. While they 'continued to be judged as reliable' under SIS post–war validation efforts, 'reports from those sources tended to present a less worrying view of Iraqi chemical and biological weapons capability than [reporting] from the sources whose reporting is now subject to doubt.'[47] No reporting explicitly discussed in the rest of the report can be directly attributed to these sources, all we know is that their information was 'less worrying'. From the context of the report's discussion, this seems likely to have been information regarding Iraq's pursuit of a 'break out capability' to quickly acquire WMD if or when the international sanctions on Iraq were to be lifted. Reports of this variety included two reports received in 2001 from two different, (at the time) recently recruited sources[48] reporting that Iraq had 'recalled its nuclear scientists in 1998[49]' to duty and one report of 'Iraqi attempts to recruit new scientists by people formerly associated with Iraq's BW programme to work on BW related research, including genetic engineering'.[50] However, in the case of the sources reporting on the nuclear scientists 'neither [was] speaking from direct, current experience'.[51]

- *Source 6* was 'new source on trial' whose information was 'received too late for inclusion in the JIC assessment [on Iraq] of 9 September, it did provide significant assurance to those drafting the government's dossier that active, current production of chemical and biological weapons was taking place. A second report from the new source, about the production of a particular chemical agent, was received later in September 2002'.[52] No further reports were received and this line of reporting was subsequently withdrawn by SIS following their post-war validation effort.[53]

In the Iraq UMINT Matrix Exercise, students are provided with a slide and/or a hand-out providing the details given above on SIS' HUMINT reporting leading up to the 9 September JIC assessment on Iraqi WMD and the September Dossier that was based largely upon that appreciation. They are exhorted to ignore post-war validation results and withdrawn reporting and concentrate on what was available in late 2002. The class is then divided into syndicates who will work together on the problem of weighing and adding up and the information from these six 'main sources' into some sort of net judgement.

**Phase I: a 'Rough and Ready' Collating Matrix**

Once the students have had reading time to work through the HUMINT information available from Butler, the first stage of the exercise is to use a 'rough and ready' matrix either on a white board or Excel spreadsheet to take what might be termed a 'straw pole' of the sources (see Figure 1). Using the 'rough and ready' matrix as a collating tool helps students to visualise the available reporting as well as do a superficial accounting task.

As can be seen from the previous section, Sources 1, 2, 3 and 6 provide reports that (on the face of it) clearly support the idea that Iraq possessed a current, in-hand WMD capability most likely in the form of chemical and biological weapons (CBW) and enhanced SCUD short-range ballistic missiles. Given the practical difficulties and lead times for scientific research and development on nuclear and biological programmes, reporting from Sources 5 and 6 might support the hypothesis of a break-out capability but not a weaponised, operational WMD capability (hence their description as 'less worrying').

This moment in the exercise has a historiographic value as well as an analytic value. As can be seen in Figure 2, at a superficial, head-count level SIS' sources in late 2002 were reporting *two to one in favour* of Iraq having an operational WMD capability. It is not hard to imagine the time-pressed Cabinet Office analyst – for whom the impending September 9 paper and evolving September Dossier are but one of a bulging in-tray of tasks – stacking up reports pro and con then going with the apparent weight of opinion as it were. Of course there were gaps in the reporting, some of it was dated and not all the sources appeared equally trustworthy. But surely describes just about any and every intelligence assessment. As Michael Hayden is famously supposed to have observed to his wife over the evening dishes during the same period: 'if it were a fact it wouldn't be intelligence'.[54]

**Phase II: Source Evaluation and Weighting**

The next stage of the exercise is essentially a Socratic dialogue with students as to whether or not they find the results of the 'rough and ready' matrix convincing. Unsurprisingly, as a general rule, they do not. What follows next is essentially a three-stage task in assigning relative weightings to the sources, and some of the problems in principle with any weighting algorithm.

*1. Review of Sources*[55]

The first stage is a critical review of the sources and the validation information about them and their reports. As most students are quick to point out, Source 2 is not *actually* a source at all but a cut-out carrying information from Source 3 (and probably other sub-agents) to his SIS controllers. Source 2 does not really warrant listing as a source at all. The next source to fall prey to critical scrutiny is Source 2. As noted above, Source 2's report of 'up to 20' Al Hussein missiles 'was subsequently

reflected in all future JIC assessments (and Government statements)' even though he was, in fact, 'passing on the comments of a sub-source, who reported only once'.[56] Thus while acting merely as courier Source 2's 'position to comment authoritatively' on his own material appears to have contaminated analytic judgements drawing sub-source whose reporting really needed to be evaluated separately. In Lord Butler's measured words, this led to the JIC giving Source 3's information 'more weight' than it 'could reasonably bear'.[57]

Thus that very sub-agent, Source 3, then comes in for considerable scrutiny and criticism with concerns raised about the actual trustworthiness of reports from a source with ulterior political motives. Source 1 then falls afoul of reporting hearsay, as potentially so do 4 and 5 who appear to report a mix of first and second hand material. The most common adjective applied to Source 6 is the British colloquialism 'dodgy' which suggests questionable trustworthiness, possibly with criminal undertones. Suggestions here vary from another opportunist dissident sympathiser to an Iraqi deception agent.

Consequently, while two thirds of agents being run by SIS in Saddam Hussein's Iraq may have proven track records of good reliability, there were very real causes for concern about the accuracy of their reporting on WMD programs. And precisely such a distinction between quality of a source and that of their individual reports is fundamental to one of the most well-established and widely adopted schemes for evaluating and grading reports, the so-called Admiralty Code.

### *2. Admiralty Code Source Grading*

The question then becomes one of how best to incorporate these judgements about the reliability of the sources and their reports into the 'rough and ready' matrix. To assign weights to the reports, the class syndicates are asked to grade the 6 'main sources' according to the current British Ministry of Defence version of the so-called Admiralty Code. In its current form, the Admiralty code generates a '6x6'[58] digraph of a letter grade from A to F representing the reliability an intelligence *source* and a numerical grade from 1 to 6 to the probable truthfulness of any specific *report* from that agent. In this formula, A represents the highest reliability, E the lowest and F 'cannot be judged' for source reliability and 1 the highest likely accuracy, 5 the lowest and 6 'cannot be judged' (see Figure 2).[59] Because the individual SIS reports are not available, students are asked to grade the sources individually but each source's overall reporting cumulatively.

Of course, in practice, a source report's grading is allocated by the originating agency such as, in the UK, SIS for national HUMINT or Government Communications Headquarters for intercept-derived

materials. However, it is worth keeping in mind that many if not most HUMINT operations are managed by teams, in the SIS case typically drawing on participation by a range of functional representatives including case officers and/or their line managers, targeting officers , operational security experts and the collating or 'requirements' officers who are immediately responsible for evaluating the reporting and circulating it to consumers, traditionally in the form of so-called CX Reports.[60] Consequently, source reports and their gradings are collective, organisational products in which an agent's hander provides but one of a number of contributing judgements. Thus this stage of the exercise also encourages students to be aware of, and reflect on, the foibles and fragilities affecting the validation judgements informing source reports. And analysts, of course, rely to no small extent on such judgements in considering how and to what degree those source reports should shape their analytic judgements.

3. *Weighted Sum Collating Matrix*

The source gradings are then compared across the syndicates and students discuss and debate their judgements supporting specific grades (which can be quite diverse at both the upper and lower ends of the scale). From this a broadly agreed 6x6 grading is then formulated, and the reporting values used to populate a weighted sum version of the 'rough and ready' matrix, with results typically resembling Figure 3. Given that low numerical values indicate higher probable reporting accuracy the rule on this version of the collating matrix is 'low score wins'. The result of this weighted sum collating matrix typically renders the net value of the available reporting in favour of Iraq *not* having WMDs.

In some respects, the 'rough and ready' matrix superficially resembles another manual matrix manipulation technique developed for analysts at the UK's Defence Intelligence Assessment Staff. The DIAS structured tradecraft manual *Quick Wins for Busy Analysts* describes a manual ACH procedure intended for whiteboard or spreadsheet. In this SAT reporting also runs down the Y axis and alternative hypothesis along the X axis with a cell for each report displaying a grading value for each report. However, the rest of the cells in the matrix are populated but bases its scoring not on the reliability of the sources rather than the likelihood of their report appearing 'if [a given] hypothesis were true'.[61] Reliability grading enters the equation only in the latter stages at which point one is advised to 'Double-check how reliable/credible the most diagnostic evidence is – especially if your conclusions hinge upon it'.[62] Thus reliability becomes a second-stage judgement while the current exercise asks, essentially, in what weight of credibility can be awarded to evidence pointing towards a particular conclusion. These are logically different questions, both of which are

analytically significant, and one of advantage of the PARC ACH tool is that it incorporates both reliability and diagnosticity considerations.

Students are then asked to scrutinize the weighted sum matrix and identify what the potential problems may be. The most significant potential problem arises, of course, from the total number of reports received where a large number of reliable reports favouring one view may push the numerical totals up distorting any result of the 'low score wins' rule. And so students are encouraged to consider possible solutions these limitations such as by using the arithmetic mean of the reporting values (and the limitations or risks of that approach). The point in this phase of the discussion is to get students to reflect on how artefacts of any weighted sum rule may affect the apparent outcome of the collating effort. Students are, in essence, asked to formulate aspects of the underpinning logic of ACH methodology for themselves.

Once the algorithmic problems of trying to collate multiple sources of varying reliability has been taken up at least superficially, module instructors are then in a position akin to the television chef who presents a copy of the featured dish that they 'made earlier'. And so PARC ACH is introduced.

**Phase III: Using PARC ACH**

The PARC ACH stage of this exercise also runs essentially in three stages. The first stage entails a re-run of the syndicate Admiralty code task but using the source evaluation used in PARC ACH. This is the standard social science combination[63] of *reliability* and *relevance*. Compared with the Admiralty Code, PARC ACH's grading scheme has much less granularity, limiting the user's judgements to 'high', 'medium' and 'low' in each category. In principle, this grading scheme would be applied to individual 'items of evidence', such as individual HUMINT reports. But the nature of the Butler Report, as noted, necessitates treating lines of reporting in their respective aggregates. Thus the ACH '3x3' grading is applied, in a sense, to each individual SIS 'main source' in Iraq, but with reference to the quality of their WMD *reporting* rather than their prior track records of reliability as *sources*. As with the Admiralty Code task, the syndicates assign their own values, and then the exercise leader draws out agreed or common working gradings through discussion with and between the teams.

At this point, we typically use a partially populated, pre-prepared PARC ACH data file (in the tool's proprietary ACHZ format). This file already has the Iraq 'main sources' listed, along with validation information about each source included in that field of the matrix (see Figure 4). How the next stage is implemented depends on on the time available for the exercise. The syndicates may set about populating the 3x3 gradings, and then consistency and inconsistency scores, per the software's built-

in instruction manual, with their final results compared in class discussion. Alternatively, and more quickly, the exercise instructor operates a single matrix on a main classroom audiovisual display, populating the matrix on the basis of 'audience participation' discussion with syndicates and individuals making their own cases for the values that ought to be assigned (see Figure 5 for a typical outcome). At this point one can use PARC ACH to see how different judgements in terms of report grading or consistency may alter the outcome, and start to feed in other, additional and more nuanced hypotheses such as that of a 'break-out capability'. These could be fed into the 'rough and ready' matrix, but it is quicker, easier and in many ways more elegant to use the software tool.

Interestingly, there is no guarantee that the software tool is any more likely to churn out the 'right' answer expected by 20:20 hindsight. While the usual result, like the 'rough and ready' matrix, is clearly in favour of 'does not have WMDs' PARC ACH is often less clear cut, and on occasion quite the opposite result can emerge. Here again, the lesson is as much historiographic as it analytical. With the validation information available at the time, the JIC's conclusions in September 2002 were far from unreasonable, as Lord Butler was often at pains to point out. The fashionably complacent and often contemptuous retrospective view of intelligence on Iraqi WMDs has always been a facile oversimplification, and few things can make students rethink that easy conventional wisdom matter better than having them work through the evidence themselves in a way that limits the influence of hindsight bias.

**Applications Beyond Iraq**

The great advantages of the Iraq case study are both that it is on a manageable data set of reporting and the range of plausible hypotheses, and the fact that it has *a known answer*. On the other hand, it is now quite a dated example nearly 15 years in the past, and, like so much intelligence literature, very much focused on the so-called Anglosphere. One has a real sense of a message successfully conveyed to see the change in expression across a classroom when students see that—when sources are scored for reliability or not—that the judgements can change significantly, even sometimes supporting the intelligence assessments of the time. Much of the CPD training conducted by BCISS lies outside the Anglosphere however, creating a demand for more current and less Anglophone examples to work from. The team have, therefore, sought new subject datasets on which to run the same exercise. In the light of concerns expressed in many quarters about possible Russian influence operations against various recent and (at the time of writing) impending elections in the West, we have recently developed a version of the exercise based on the January 2017 'Trump-Russia' dossier, produced by former SIS officer Christopher Steele. Examining evidence of Russian interference in the US 2016 Presidential election, Steele originally produced the dossier for Republican Party

opponents of Trump with payment for Steele's services later taken over by Democrat sources. He eventually produced it *pro bono* and it was was released initially by *Buzzfeed* but hinted at by other media between October 2016 and its release.[64] Unlike Iraq, the outcome of this assessment remains keenly contested, with President Trump calling it "fake news"[65] but other sources viewing it as potentially highly credible.[66]

To begin, exercising teams can generate robust competing hypotheses, ranging from "There were no links between Trump and Russia" to "Trump is subject to Russian *kompromat"* with one or more intermediate options. Against these hypotheses the component sources reported in the dossier can be loaded. Steele's dossier includes 17 different reports, which is too many to be manageable in the short exercise described here. However, the first report of the dossier, the two-page "Company Intelligence Report 2016/080", breaks down its assertions by letter-coded sources (A-E, G) allowing a more limited scope of information. Since the Dossier does not grade its sources, the exercise can then proceed to weight the manual collating matrix and PARC ACH table with credibility scores which the students can glean from the short descriptions provided in the dossier. More ambitious exercises could add further individual reports from within the dossier. While not providing the fidelity offered by the 2003 WMD dataset, there is immediacy to the Trump-Russia exercise, a direct relevance to experiences beyond the United States as well as sufficient complexity and ambiguity to make it a worthwhile version of the exercise. It should be possible to run versions of this exercise with a wide range of different scenarios – both positive and counterintelligence – where commensurable information sets are available.

**Conclusion**

PARC ACH is, of course, over a decade old now and is no longer maintained by the team at Palo Alto Research Centre. Occasionally backwards compatibility problems present themselves, especially when working cross-platform on the supposedly platform-independent Java implementations. Certain features, such as matrix duplication, have ceased to work, but one can work around this by opening multiple instances of the application and printing ACHZ tables has always been problematic. It remains, however, the most cost-efficient solution available for this kind of teaching, supported by an excellent built-in operating manual and good conceptual and technical support in the professional and academic literature on analytic tradecraft. Whatever its value to working analysts might be or have been, it remains without equal for teaching core concepts in intelligence studies to do with source evaluation, the potential impact of such validation on analytical judgements, and the larger challenges of fusing significant volumes of diverse information in the intelligence assessment

process. Scholars and practitioners alike can benefit from this tool (and tools like it) for very different reasons, especially when it is part of a wider pedagogical strategy.

Like many of the practicals in the BCISS training cum education repertoire (as in many intelligence-oriented teaching programmes elsewhere), the goal of the Iraq HUMINT Matrix exercise is to get participants to think through and work out analytic methods, issues and potential solutions from first principles and for themselves. The pedagogic goal is to try and fuse training and education learning outcomes, so that students emerge with a technical competence in analytic methods underpinned by a deep understanding of the foundations and internal logic shaping those methods. The Iraq Matrix exercise seeks to unpack and examine the nuts and bolts of source evaluation and testing alternative hypotheses with particular attention to the relationship between the quality of ones sources and, *après* Lord Butler, the weight of judgements they can or cannot sustain. It is a running joke amongst BCISS students that almost nothing taught on the Centre's programmes ever actually makes their lives or work *easier*. The ultimate goal is of that teaching is to encourage what is currently fashionably referred to as 'reflexive practice', where the practitioner reflects critically and self-critically upon how their task works and how they do it and uses those insights to improve their workplace performance. But as we suggested at the beginning of this discussion, not all of our teaching is directed towards *practitioners*. For those whose aims are scholarly and academic, the goal of an exercise like this is to give observers a more visceral understanding of the challenges of the intelligence task they intended to study. And here the intended reflexive practice goal is to encourage an empathy with the workaday challenges facing those and to discourage the natural observer's temptation to make facile and simplistic judgements about processes or events.

**Diagrams**

| Sources/Hypotheses | Iraq Has CBW | Iraq Does Not Have CBW |
|---|---|---|
| 1 | ✓ | |
| 2 | ✓ | |
| 3 | ✓ | |
| 4 | | ✓ |
| 5 | | ✓ |
| 6 | ✓ | |
| **Net Judgement** | **4** | **2** |

Figure 1: 'Rough and Ready' Collating Matrix

| Reliability of the Source | | Credibility of the Information | |
|---|---|---|---|
| **A** | Completely reliable | **1** | Confirmed by other sources |
| **B** | Usually reliable | **2** | Probably true |
| **C** | Fairly reliable | **3** | Possibly true |
| **D** | Not usually reliable | **4** | Doubtful |
| **E** | Unreliable | **5** | Improbable |
| **F** | Reliability cannot be judged | **6** | Truth cannot be judged |

Figure 2: Current UK 'Admiralty Code' (JDP 2-00)

| Sources /Hypotheses | Iraq Has CBW | Iraq Does Not Have CBW |
|---|---|---|
| 1 | 3 | |
| 2 | ---- | --- |
| 3 | 4 | |
| 4 | | 2 |
| 5 | | 2 |
| 6 | 6 | |
| **Net Judgement** | **13** | **4** |

Figure 3: Weighted Sum 'Low Score Wins' 'Rough and Ready' Matrix

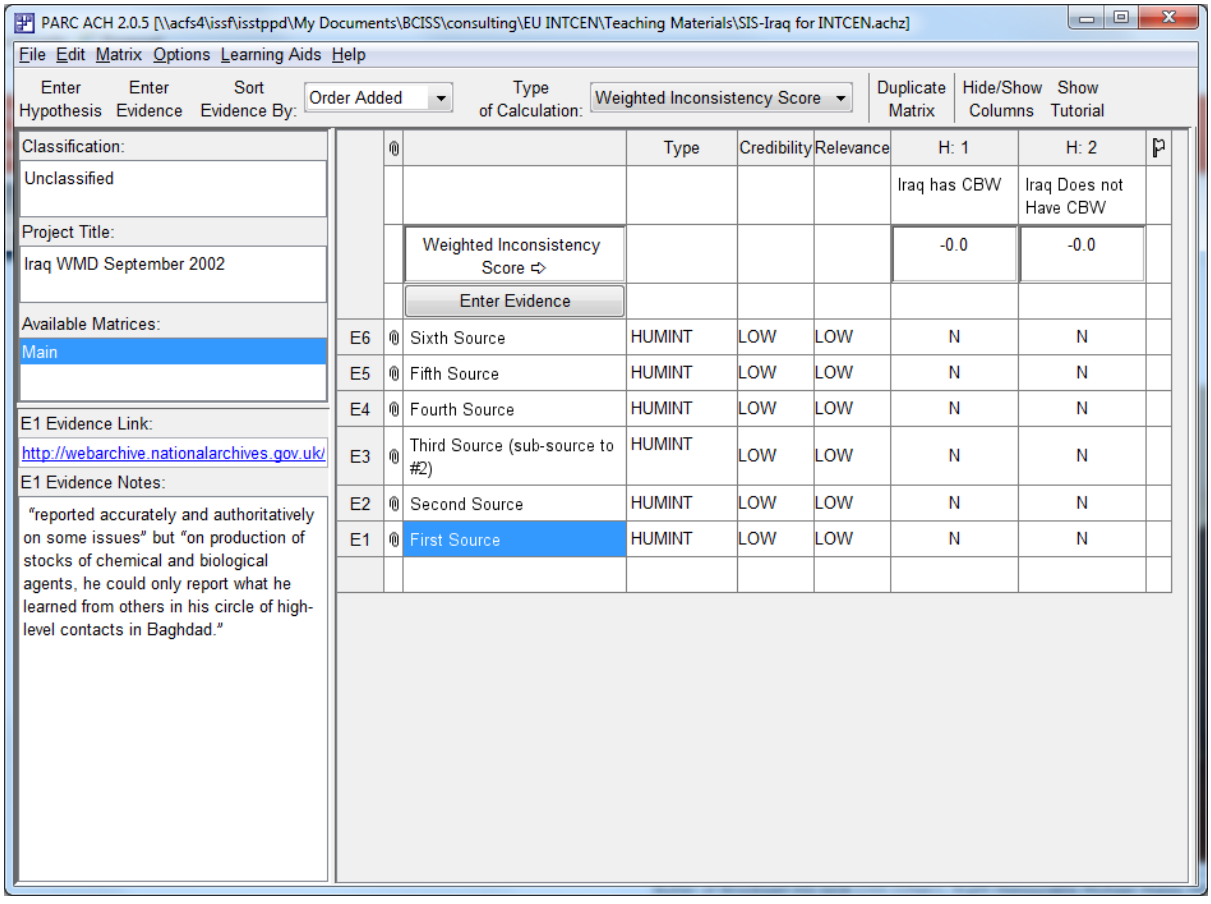


Figure 4: Pre-Prepared Iraq HUMINT Matrix PARC ACH File – with dummy values.

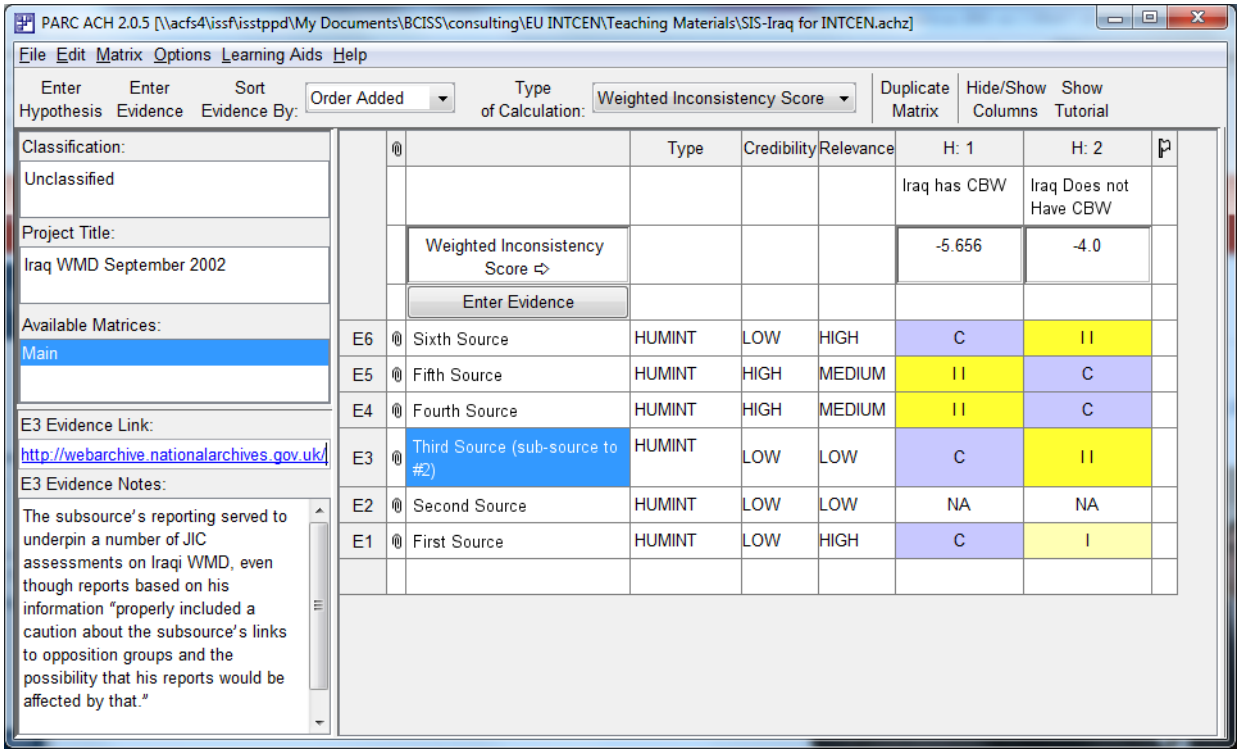


Figure 5: Indicative Completed PARC ACH Iraq HUMINT Matrix – on this occasion the 'Right Answer'

**References**


Robert David Booth *State Department Counterintelligence: Spies, Leaks and Lies* (Dallas, TX: Brown Books Publishing Group, 2014)

Bensinger, Ken, Miriam Elder and Mark Schoofs, "These Reports Allege Trump Has Deep Ties To Russia", *Buzzfeed*, 10 Jan 2017, https://www.buzzfeed.com/kenbensinger/these-reports-allege-trump-has-deep-ties-to-russia?utm_term=.jheD5OmKJ#.htWO4Y23q, (accessed 15 April 2017).

Butler, Lord Robin. *Review of Intelligence on Weapons of Mass Destruction* (London: The Stationery Office [TSO], 2004).

Chilcot, John. *Report of the Iraq Inquiry* (London: Willams Lea Group, 2016).

Clausewitz, Carl von *On War* (Howard & Paret, trans) (Princeton NJ: Princeton University Press, 1989)

Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction (aka WMD Commission) *Report to the President of the United States* (Washington DC: United States Government Printing Office, 2005)

Dahl, Erik *Intelligence and Surprise Attack: Failure and Success from Pearl Harbor to 9/11 and Beyond* (Washington, DC: Georgetown University Press, 2013)

Davies, Philip H.J. 'Assessment BASE: Simulating National Intelligence Assessment in a Graduate Course' *International Journal of Intelligence and CounterIntelligence* 19:4 (October 2006): 721-736

Davies, Philip H.J. 'Collection and Analysis on Iraq: a Critical Look at Britain's Spy Machinery' *Studies in Intelligence* Vol.49 No.4 (October 2005): 41-54.

Davies, Philip H.J. 'Jointery Versus Tradecraft', William Lahneman and Ruben Arcos ed. *The Art of Intelligence: Simulations, Exercises and Games* (New York: Rowman and Littlefield, 2014): 203-222.

Davies, . Philip H.J. *MI6 and the Machinery of Spying* (London: Frank Cass, 2004).

Development, Concepts and Doctrine Centre (DCDC). *JDP 2-00: Understanding and Intelligence Support to Joint Operations* 3rd Edition (Shrivenham, UK: DCDC, 2011).

Dylan, Huw. 'Thinking About Defence Intelligence: Victor Cavendish-Bentinck, Denis Capel-Dunn, Kenneth Strong and the Joint Intelligence Bureau as Foundation for the Defence Intelligence Staff' *Intelligence and National Security* Vol.31 No.6 (October 2016): 829-843.

Henderson, Barney , David Lawler and Louise Burke. 'Donald Trump attacks alleged Russian dossier as "fake news" and slams Buzzfeed and CNN at press conference'. *The Telegraph,* 17 Jan 2017. Retrieved 15 April, 2017.

Heuer, Richards. *The Psychology of Intelligence Analysis* (Washington DC: Center for Intelligence Studies, 1999)

Heuer, Richards and Randolph H. Pherson. *Structured Analytic Techniques for Intelligence Analysis* (Washington DC: CQPress, 2011).

House of Commons Foreign Affairs Select Committee. *The Decision to Go to War in IraqHC 813-I* (London: The Stationery Office, 2003)

Hutton, Lord Brian. *Report of the Inquiry into the Circumstances Surrounding the Death of Dr David Kelly C.M.G* (London: TSO, 2004).

i2 Homepage at https://www.ibm.com/us-en/marketplace/enterprise-intelligence-analysis (accessed 17 April 2017).

Intelligence and Security Committee. *Iraqi Weapons of Mass Destruction: Intelligence and Assessments* (London: TSO, 2003).

Joint Chiefs of Staff *JP 2-01 Joint and National Intelligence Support to Military Operations* (Washington DC: Department of Defence, 2012)

Jomini, Antoine Henri. *Art of War* (Mendell & Craighill, trans). (Rockville, MD: Arc Manor, 2007)

MacEachin, Douglas. 'The Tradecraft of Analysis' in Roy Godfson, Ernest R. May and Gary Schmitt ed. *US Intelligence at the Crossroads: Agendas for Reform* (Washington DC: Brassey's, 1995) pp.63-74.

Paterva Maltego Homepage https://www.paterva.com/web7/ (accessed 17 April 2017).

Marrin, Stephen *Improving Intelligence Analysis: Bridging the Gap Between Scholarship and Practice* (London: Routlege, 2011).

Mencken, Henry Louis. *Prejudices: Second Series* (New York: Alfred A. Knopf, 1920)

Palacios, José-Miguel. 'Intelligence Analysis Training: A European Perspective' *International Journal of Intelligence, Security and Public Affairs* Vol.18 No.1 (2016): 34-56.

Palantir Homepage https://www.palantir.com/ (accessed 17 April 2017).

Palo Alto Research Centre. '$ACH_{2.0.5}$ Download Page' http://www2.parc.com/istl/projects/ach/ach.html (accessed 14 April 2017).

Professional Head of Defence Intelligence Analysis (PODIA) *Quick Wins for Busy Analysts* (London: Ministry of Defence, 2012).

Steele, Christopher. 'The Trump Russia Dossier', https://www.documentcloud.org/documents/3259984-Trump-Intelligence-Allegations.html, (accessed 15 April 2017)

Stober, Dan and Ian Hoffman *A Convenient Spy: Wen Ho Lee and the Politics of Nuclear Espionage* (London: Simon & Schuster, 2001)

Strong, Kenneth *Men of Intelligence* (London: Cassell, 1970)

Tracy, Abigail. "What Intelligence Experts Think of the Explosive Trump-Russia Report", *Vanity Fair* (January 11, 2017).

United States Army *FM 34-60 Counterintelligence* (Washington DC: Headquarters of the Army, 1995)

US Government. *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (Washington DC: Centre for the Study of Intelligence, 2009)

Woodward, Bob. *Plan of Attack* (New York: Simon & Schuster, 2004).

[1] Author for correspondence purposes.
[2] Henry Louis Mencken *Prejudices: Second Series* (New York: Alfred A. Knopf, 1920) p.120.
[3] Philip H.J. Davies 'Assessment BASE: Simulating National Intelligence Assessment in a Graduate Course' *International Journal of Intelligence and CounterIntelligence* 19:4 (October 2006) pp.721-736 and 'Jointery Versus Tradecraft', William Lahneman and Ruben Arcos ed. *The Art of Intelligence: Simulations, Exercises and Games* (New York: Rowman and Littlefield, 2014) pp.203-222.
[4] For an independent account, see José-Miguel Palacios 'Intelligence Analysis Training: A European Perspective' *International Journal of Intelligence, Security and Public Affairs* Vol.18 No.1 (2016) p.48.
[5] Stephen Marrin *Improving Intelligence Analysis: Bridging the Gap Between Scholarship and Practice* (London: Routlege, 2011) p.77.
[6] (London: The Stationery Office (TSO), 2004), hereafter just referred to as the *Review*.
[7] This is a recurrent theme throughout report of the Presidential Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction (aka WMD Commission) *Report to the President of the United States* (Washington DC: United States Government Printing Office, 2005) but is particularly explicit on p.169.
[8] Palo Alto Research Centre. '$ACH_{2.0.5}$ Download Page' http://www2.parc.com/istl/projects/ach/ach.html (accessed 14 April 2017).
[9] Since acquired by IBM as i2 Analyze, see https://www.ibm.com/us-en/marketplace/enterprise-intelligence-analysis (accessed 17 April 2017).
[10] https://www.palantir.com/ (accessed 17 April 2017).
[11] https://www.paterva.com/web7/ (accessed 17 April 2017).
[12] For which the fiueld has been sternly criticised by, e.g. Erik Dahl in his *Intelligence and Surprise Attack: Failure and Success from Pearl Harbor to 9/11 and Beyond* (Washington, DC: Georgetown University Press, 2013).
[13] Philip H.J. Davies 'Collection and Analysis on Iraq: a Critical Look at Britain's Spy Machinery' *Studies in Intelligence* Vol.49 No.4 (October 2005) pp.41-54.
[14] US Government. *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis* (Washington DC: Centre for the Study of Intelligence, 2009) pp.14-15.
[15] For an overview of Strong's background and immediate post-war intelligence career see Huw Dylan 'Thinking About Defence Intelligence: Victor Cavendish-Bentinck, Denis Capel-Dunn, Kenneth Strong and the Joint Intelligence Bureau as Foundation for the Defence Intelligence Staff' *Intelligence and National Security* Vol.31 No.6 (October 2016) esp. pp.836-843.
[16] Kenneth Strong *Men of Intelligence* (London: Cassell, 1970) p.40. The authors are indebted to Matt Jolly and to a former senior official at DIAS for bringing Strong's account to our attention.
[17] See, e.g. Carl von Clausewitz, *On War* (Howard & Paret, trans) (Princeton NJ: Princeton University Press, 1989) p. 117.
[18] Antoine Henri de Jomini, *Art of War* (Mendell & Craighill, trans). (Rockville, MD: Arc Manor, 2007) p. 203.
[19] Jomini, *Art of War* p.201
[20] Kenneth Strong *Men of Intelligence* p.153.
[21] Strong *Men of Intelligence* p.40 *infra*.
[22] Richards Heuer and Randolph H. Pherson *Structured Analytic Techniques for Intelligence Analysis* (Washington DC: CQPress, 2011) pp.140-143.
[23] A point Heuer and Pherson themselves note *Structured Analytic Techniques* p.141.
[24] See, e.g. Douglas MacEachin. 'The Tradecraft of Analysis' in Roy Godfson, Ernest R. May and Gary Schmitt ed. *US Intelligence at the Crossroads: Agendas for Reform* (Washington DC: Brassey's, 1995) pp.63-74.

[25] See, e.g. United States Army *FM 34-60 Counterintelligence* (Washington DC: Headquarters of the Army, 1995) pp.45-48 or more recently Joint Chiefs of Staff *JP 2-01 Joint and National Intelligence Support to Military Operations* (Washington DC: Department of Defence, 2012) pp.D9 – D13.
[26] Dan Stober and Ian Hoffman *A Convenient Spy: Wen Ho Lee and the Politics of Nuclear Espionage* (London: Simon & Schuster, 2001) p.235; significantly this method failed in the investigation of Chinese acquisition of details concerning the W88 thermonuclear warhead largely because of the large number of individuals with access to information on the programme.
[27] Robert David Booth *State Department Counterintelligence: Spies, Leaks and Lies* (Dallas, TX: Brown Books Publishing Group, 2014) *passim*.
[28] For example, Richards Heuer *The Psychology of Intelligence Analysis* (Washington DC: Center for Intelligence Studies, 1999) p.19
[29] House of Commons Foreign Affairs Select Committee *The Decision to Go to War in IraqHC 813-I* (London: The Stationery Office 2003) esp. pp.48-50.
[30] Intelligence and Security Committee *Iraqi Weapons of Mass Destruction: Intelligence and Assessments* (London: TSO, 2003); on partisan political influence especially on the then-Chair, Anne Tayler MP see e.g. Anthony Glees and Philip H.J. Davies Intelligence, Iraq and the Limits of Legislative Accountability during Political Crisis *Intelligence and National Security* Vol.21 No.5 (October 2006) pp.848-883.
[31] Lord Brian Hutton *Report of the Inquiry into the Circumstances Surrounding the Death of Dr David Kelly C.M.G* (London: TSO, 2004).
[32] Butler *Review*
[33] 'Whitehall Mandarin' is a colloquialism of the British political classes denoting a certain kind of cautiously phrased, clinical and sometimes technocratically opaque style of drafting considered typical of the Home Civil Service.
[34] Fortunately also available as a pdf on an open license; Sir John Chilcot *Report of the Iraq Inquiry* (London: Willams Lea Group, 2016) downloadable pdf at: http://www.iraqinquiry.org.uk/the-report/ (accessed 17 April 2017).
[35] Considerably more detail was, however, given to the finished JIC assessments.
[36] As Butler notes at a number of points the *Review*, the lion's share of UK assessments on Iraq WMD dealt with and pointed towards a 'break out capability' more than anything else pp.41, 48, 63, 75, 116, 130, 157 and *passim*.
[37] This appears to have resulted from last minute edits trying to present SIS sources in a manner that sufficiently protected their identity. Private information.
[38] Butler *Review* pp.100-101 and 107-108 respectively.,
[39] Note that the order in which these sources are treated follows that in Butler's *Review* and not that used in Davies 'Collection and Analysis on Iraq' pp.46-58.
[40] Butler *Review* pp.100; 107.
[41] Butler *Review* p.75.
[42] Butler *Review* pp.100, 107.
[43] Butler *Review* p.100.
[44] Butler *Review* p.107.
[45] Private information.
[46] Butler *Review* p.61.
[47] Butler Review pp.100, 102, 108.
[48] Another confusing feature of the Butler report is the appearance of sources described as shown to be reliable in 2004 appearing as 'new' sources at the times of their recruitment two or three years previously.
[49] In the phasing used in a JIC assessment of 10 May 2001, Butler *Review* p.56.
[50] Butler *Review* p.60.
[51] Butler *Review* p.60 *infra*.
[52] Butler *Review* p.100-101.
[53] Butler *Review* p.101.
[54] Bob Woodward, *Plan of Attack* (New York: Simon & Schuster, 2004), 219.
[55] Referred to in the CIA *Primer* as a 'quality of information check' pp.10-11.
[56] Butler *Review* p.61.
[57] Butler *Review* p.56.
[58] Much more commonly seen is the older '5x5' grading convention.

[59] Development, Concepts and Doctrine Centre (DCDC) *JDP 2-00: Understanding and Intelligence Support to Joint Operations* (Shrivenham, UK: DCDC, 2011).
[60] See, e.g. Philip H.J. Davies *MI6 and the Machinery of Spying* (London: Frank Cass, 2004) pp.287, 338-339.
[61]Professional Head of Defence Intelligence Analysis (PODIA) *Quick Wins for Busy Analysts* (London: Ministry of Defence, 2012) p.15.
[62] PODIA *Quick Wins* p.17
[63] In most implementations of this practical by BCISS the matrix exercise is preceeded by at least one lecture on the fundamentals of social science research concepts and methodology.
[64] Christopher Steele, 'The Trump Russia Dossier', available at https://www.documentcloud.org/documents/3259984-Trump-Intelligence-Allegations.html, accessed 15 April 2017. For the original publication, see Ken Bensinger, Miriam Elder and Mark Schoofs, 'These Reports Allege Trump Has Deep Ties To Russia', *Buzzfeed*, 10 Jan 2017, https://www.buzzfeed.com/kenbensinger/these-reports-allege-trump-has-deep-ties-to-russia?utm_term=.jheD5OmKJ#.htWO4Y23q, (accessed 15 April 2017).
[65] Barney Henderson, David Lawler and Louise Burke 'Donald Trump attacks alleged Russian dossier as "fake news" and slams Buzzfeed and CNN at press conference'. *The Telegraph,* 17 Jan 2017. Retrieved 15 April, 2017.
[66] Abigail Tracy, "What Intelligence Experts Think of the Explosive Trump-Russia Report", *Vanity Fair* (January 11, 2017).