

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Cognitive Biases in a Geospatial Intelligence Analysis Task: An ACT-R Model

Permalink

<https://escholarship.org/uc/item/9c97q1b4>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN

1069-7977

Authors

Paik, Jaehyon
Pirolli, Peter
Lebiere, Christian
et al.

Publication Date

2012

Peer reviewed

Cognitive Biases in a Geospatial Intelligence Analysis Task: An ACT-R Model

Jaehyon Paik (jpaik@parc.com), Peter L. Pirolli (pirolli@parc.com)

Palo Alto Research Center
3333 Coyote Hill Rd., Palo Alto, CA 94304

Christian Lebiere (cl@andrew.cmu.edu), Matthew Rutledge-Taylor (mattrt@andrew.cmu.edu)

Carnegie Mellon University, Psychology Department
5000 Forbes Avenue, Pittsburgh, PA 15218

Abstract

An ACT-R model of sensemaking in a geospatial intelligence task was developed based on Instance-Based Learning Theory (IBLT). The model (a) maintains hypotheses about the probability of attacks by insurgent groups, (b) seeks new information based on those hypotheses, and (c) updates hypotheses based on new evidence. The model provides a functional account of how these sensemaking processes are carried out in a cognitive architecture, and model performance can be compared to normative (Bayesian) standards. Simulations exhibit two well-known cognitive biases that are frequently identified as problems in intelligence analysis: (1) anchoring in the weighting of new evidence and (2) confirmation bias in seeking new information.

Keywords: ACT-R, cognitive biases, sensemaking

Introduction

Sensemaking (Klein, Moon, & Hoffman, 2006a, 2006b; Pirolli & Card, 2005; Russell, Stefik, Pirolli, & Card, 1993) is a concept that has been used frequently in studies of intelligence analysis. The term suggests an active seeking and processing of information to achieve understanding. Sensemaking involves a set of processes aimed at seeking and filtering information, plus a set of processes that develop representational schemas (frames) that best fit the available evidence and provide a basis for understanding the data. In this paper we present the cognitive model of basic sensemaking processes for an intelligence analysis task. A major concern in the intelligence community is the impact of cognitive biases on the accuracy of analyses (Heuer, 1999). We present simulation results that exhibit *anchoring bias* in the evaluation of new evidence and *confirmation bias* in seeking evidence.

The Geospatial Task

The geospatial task (Figure 1) is one of a set of challenge tasks developed as part of the IARPA ICARUS program to drive the development of integrated neurocognitive models of sensemaking. This specific task required reasoning based on a set of rules concerning the relation of observed evidence to the likelihood of attack by four different groups. A layered geospatial map is presented on a computer screen, with different layers presenting different forms of intelligence (INTs). The INTs include HUMINT (human intelligence), IMINT (image intelligence), MOVINT (movement intelligence), SIGINT (signal intelligence),

SOCINT (socio-cultural intelligence), and SIGACT (attack intelligence).

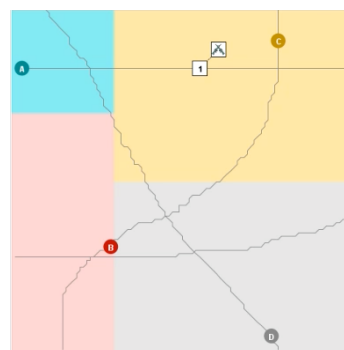


Figure 1: The screen shot of the geospatial task. The letters (A, B, C, D) indicate the center of the group location, and '1' surrounded by a box indicates the attack location.

The task begins with a given attack location (SIGACT) along with group centers (HUMINT), A, B, C, and D representing the center of activity for four possible insurgent groups. The first step is to report probabilities of attack by each group [A%, B%, C%, D%] based on the SIGACT and HUMINT (see Table 1)¹. After that, the task is to iteratively choose among the four remaining INT layers (Table 1), up to a total of three INTs (layers), one at a time, in any order. Each INT layer provides unique evidence. Specifically, IMINT can reveal whether an attack happened on a government or military building, MOVINT provides evidence whether an attack occurred in dense or sparse traffic, SIGINT indicates electronic “chatter” or “silence” by different groups, and SOCINT indicates the group whose region the attack happened. At each stage, the selection of a particular INT provides evidence that can be used to update the probability distribution over the hypotheses about the responsibility of the four groups in producing the given attack. The rules specifying how evidence ought to update these probabilities is given in the PROBS rules in Table 1. After the last stage of INT selection, the task is to allocate resources (troops) to prevent further attacks.

¹ The new version of the task will provide the initial probabilities based on HUMINT

Table 1: Probabilistic rules provided to user for inferring beliefs about group attack likelihoods.

INTS	PROBS
HUMINT	If a group attacks, then the relative likelihood of attack decreases as the distance from the group center increases.
IMINT	If A or B attack then the attack is four times as likely to occur on a <i>Government</i> versus <i>Military</i> building. If C or D attack then vice versa.
MOVINT	If A or C attack then the attack is four times as likely to occur in <i>dense</i> versus <i>sparse</i> traffic. If B or D attack then vice versa.
SIGINT	If SIGINT on a group reports <i>chatter</i> , then attack by that group is seven times as likely as attack by each other group If SIGINT on a group reports <i>silence</i> , then attack by that group is one-third as likely as attack by each other group.
SOCINT	If a group attacks then that group is twice as likely to attack in its own versus other region.

Anchoring and Confirmation Biases

Anchoring and confirmation biases have a long history of study in cognitive psychology and the intelligence communities (Heuer Jr, 1999; Klayman, 1995; Klayman & Ha, 1987; Nickerson, 1998; Tversky & Kahneman, 1974; Wason, 1960). Process models of these biases, especially in complex tasks, remain largely unexplored. In this paper we develop cognitively plausible process model of the geospatial task in the ACT-R architecture. We then compare this ACT-R model against a rational Bayesian model of the task to examine evidence of anchoring and confirmation biases.

Anchoring Bias and Anchoring and Adjustment Heuristic

Anchoring is a cognitive bias that occurs when individuals establish some belief based on some initial evidence, and then overly rely on this initial decision in their weighting of new evidence (Tversky & Kahneman, 1974). Human beings tend to anchor on some estimate or hypothesis and subsequent estimates tend to be adjustments that are influenced by the initial anchor point—they tend to behave as if they have an *anchoring+adjustment heuristic*. Adjustments tend to be insufficient in the sense that they overweight the initial estimates and underweight new evidence.

Confirmation Bias

Confirmation bias is typically defined as (for a survey, see Nickerson, 1998):

- The interpretation of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand (Nickerson, 1998)
- The tendency for people to seek information and cues that confirm the tentatively held hypothesis or belief,

and not seek (or discount) those that support an opposite conclusion or belief (Wickens & Hollands, 2000). The seeking of information considered supportive of favored beliefs (Nickerson, 1998).

Studies (Cheikes, Brown, Lehner, & Adelman, 2004; Convertino, Billman, Pirolli, Massar, & Shrager, 2008; Tolcott, Marvin, & Lehner, 1989) have found evidence of confirmation bias in tasks involving intelligence analysis, and there is a common assumption that many intelligence failures are the result of confirmation bias in particular (Chorev, 1996; Grabo & Goldman, 2004; Heuer Jr, 1999).

Biases in the Geospatial Task

The geospatial task might elicit anchoring and confirmation biases at multiple points in the process. Anchoring bias in weighing evidence might be found when participants revise their belief probabilities after selecting and interpreting a particular INT. The estimates of belief probabilities that were set prior to the new INT evidence could act as an anchor, and the revised (posterior) belief probabilities could be insufficiently adjusted to reflect the new INT (i.e., when compared to some normative standard).

Confirmation bias in weighing evidence can also be found in the hypothesis adjustment process. When applying a particular INT, such as IMINT (which supports multiple hypotheses), participants may only apply the adjustment to the preferred hypothesis while neglecting other groups also supported by evidence, or weight the evidence too strongly in favor of the preferred hypothesis.

Finding confirmation bias in seeking evidence in the task is somewhat more difficult since most INTS apply equally to all hypotheses. We used the SIGINT layer to identify this kind of bias because a single hypothesis has to be selected for that layer. SIGINT provides considerable gains to the selected hypothesis when chatter is detected (7 times more likely), so participants could get significant certainty. However, it loses considerable weight (3 times less likely) when silence is detected. Thus, a decision to choose the SIGINT layer too early (before a specific group has dominated the other in terms of relative likelihood) might be interpreted as confirmation bias in evidence seeking.

The ACT-R architecture

ACT-R (Anderson et al., 2004; Anderson & Lebiere, 1998) is a cognitive architecture that includes a declarative memory module that stores and retrieves information and a procedural module that coordinates the flow of information. Declarative knowledge in ACT-R is represented formally as *chunks* of information (Miller, 1956; Simon, 1974). Chunks are recalled from long-term declarative memory by an activation based retrieval process. Activation spreads from the current focus of attention, including goals, through *associations* among chunks in declarative memory. The spread of activation from one cognitive structure to another is determined by attentional weights on the associations among chunks. These weights determine the rate of activation flow among chunks. *Partial matching* is a

mechanism that allows for chunks in declarative memory that do not perfectly match a retrieval request to be retrieved. *Blending* is a memory retrieval mechanism that allows all chunks in declarative memory that match or partially match a retrieval request to blend together to create a new chunk representing an aggregate response (Lebiere, 1999).

Production rules are used to represent procedural knowledge in ACT-R. That is, they specify how to apply cognitive skill (know-how) in the current context, and how to retrieve and modify information in other modules. In ACT-R, each production rule has conditions that specify structures that are matched in limited-capacity buffers corresponding to information from the external world or other internal modules. Each production rule has actions that specify changes to be made to the buffers or requested functions in the associated modules.

The Rational Model versus ACT-R Model

We developed an ACT-R model to perform the geospatial task, as well as a rational (Bayesian) model as a normative benchmark. The ACT-R model implemented a version of instance-based learning theory (Gonzalez, Lerch, & Lebiere, 2003), and the rational model employs a standard Bayesian approach for updating belief and selecting new evidence (INTs) based on an expected information gain metric.

The Rational Model

From the PROBS rules discussed in table 1, we can extract specifications of the likelihoods of evidence, $P(e|h)$, where e is evidence (e.g., “chatter”) and h is a hypothesis (“group A attacks”). Bayes rule can be applied to compute the posterior likelihood of h given specific evidence e and prior probabilities $P(h)$

$$P(h|e) = \frac{P(e|h)P(h)}{\sum_i P(e|i)P(i)}$$

where i iterates over all hypotheses.

For instance, in Figure 2, we assume some HUMINT data has been processed, a probability has been assigned to each of the hypotheses, and the goal is to evaluate the choice of an IMINT layer. The outcomes represent the estimates of government and military building attacks given the current hypotheses strengths. The posteriors are the updated probability distributions according to the outcomes.

The choice of INT layers can be evaluated by their effects on *expected information gain* (Austerweil & Griffiths, 2011). Information gain is defined as the reduction in entropy measured over the hypothesis probabilities that occur by acquiring additional evidence. Information gain is specified as

$$IG(D, e) = H(D) - H(D|e)$$

where $H(D)$ is the entropy of the distribution of probabilities over hypotheses, and $H(D|e)$ is the entropy of the distribution of posterior probabilities after some evidence e has been discovered.

$$H(D|e) = - \sum_i P(D|e) \log_2 P(D|e)$$

In Figure 2, the information gain for seeing an attack on a government building is .4 and an attack on a military building is .09. The *expected information gain* is calculated by weighting each of the possible outcomes of information gain by the probability of obtaining that outcome. Thus, the expected information gain for selecting the IMINT layer is $(.56)(.4) + (.44)(.09) = .26$

Our rational model computed the expected information gain for all layers at each stage. The rational choices were compared to the selections made by ACT-R to identify biases.

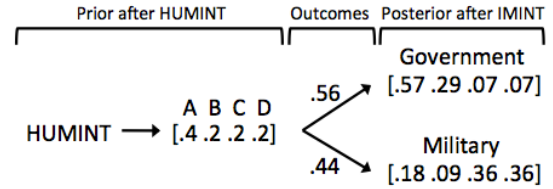


Figure 2: An example of the rational Bayesian hypothesis estimates for an IMINT layer selection.

The ACT-R Model

We assume that an average person is not able to compute the expected information gain of all possible layers, because it involves substantial amounts of computation. We considered two cognitively plausible alternatives to develop an ACT-R model.

- *Difference reduction heuristics.* One cognitively plausible way to reduce complexity is to assume that people use a heuristic such as hill climbing to evaluate moves. Rather than focus on maximizing expected information gain, hill-climbing analysis could focus on achieving states that are closer to an ideal goal state (i.e., in this case, a state in which the attacks are unambiguously caused by Group A, or Group B, etc.). This would require some heuristic for evaluating differences (distances) from the goal state.
- *Memory-based move evaluation.* It is well known in the field of naturalistic decision making that experts invariably rely on vast amounts of declarative memory experience and well-practiced cognitive skill (Klein, 1998). We assume that participants store move outcomes in declarative memory, and that blended retrievals based on current states and possible moves can produce a blended retrieval of outcomes to those moves. This would be a weighted smoothing of gains that had been made by similar moves in the past. Although not precisely equivalent to the computation of rational expected information gains (a weighting over the gains achieved by possible layer selection outcomes), blending over memory of past INT outcomes and gains should produce similar effects.

Our ACT-R model of the geospatial task explored some plausible difference reduction heuristics in a memory-based move evaluation framework. The following weighted distance function assumes that the goal is to achieve certainty on one of the hypotheses (i.e., $p_i = 1$).

$$\sum_{i \in \text{hypotheses}} p_i(1 - p_i)$$

We assume that the model relies on the use of declarative chunks that represent hypothetical past experiences of selecting INT layers. This is intended to capture a hypothesized learning process whereby participants have attended to a current probability distribution, chosen a layer, revised their estimates of the hypotheses, and assessed the utility of the layer selection they just made. For instance, if a participant had experienced two situations in which they had assessed a probability distribution [.4 .2 .2 .2] and selected an IMINT layer, and had experienced a “government building” attack one time and a “military building” attack a second time (See figure 2). The model assumes the two chunks in its declarative memory.

(exp1	(exp2
isa layer-choice	isa layer-choice
prior-a 0.4	prior-a 0.4
prior-b 0.2	prior-b 0.2
prior-c 0.2	prior-c 0.2
prior-d 0.2	prior-d 0.2
layer IMINT	layer IMINT
outcomes government	outcomes military
utility 0.58)	utility 0.69)

where the utilities are computed by the weighted distance metric.

At a future layer selection point, a production rule will request a blended/partial matching retrieval from declarative memory like below:

```
+blending>
  isa layer-choice
  prior-a 0.45
  prior-b 0.15
  prior-c 0.15
  prior-d 0.25
  layer IMINT
  utility =utility
```

This retrieval will partially match against the experience chunks above, and will blend across the stored utilities for all experienced IMINT outcomes (i.e., both government and military building experiences in the past) to produce a kind of “expected” utility to match the =utility request.

Hypothesis Probability Updating

Lebiere (1999) proposed a model of cognitive arithmetic that used retrieval of arithmetic facts to generate estimates of answers without explicit computations. The cognitive arithmetic model uses partial matching to retrieve facts related to the problem, and uses the blending mechanism to merge them together to issue an aggregate estimated answer. The model reproduced a number of characteristics of the distribution of errors in elementary school children, including both table and non-table errors, error gradients around the correct answer, higher correct percentage for tie

problems, and, most relevant here, a skew toward underestimating answers, as is common in anchoring and adjustment processes.

This approach was leveraged in the current model to account for how the PROBS rules (from table 1) are interpreted and applied to estimate the effects of the rules on the relative probabilities that the groups are responsible for the attack under examination. The ACT-R model’s memory was populated with a range of facts consisting of triplets: an initial probability, an adjustment factor, and the resulting probability. These chunks are derived from the PROBS rules shown in Table 1. For example, if the attack is found to occur on of road with dense traffic, the MOVINT rule specifies that groups A and C are 4 times as likely to have been responsible. When a layer of information is made available to the model, it adjusts the current set of probabilities by retrieving the relevant chunks and replacing the prior probabilities with the posteriors representing in the retrieved chunks. The results of this chunk based rule interpretation were then averaged over a thousand runs, given the variations in answers resulting from activation noise in the retrieval process. When provided with ratio similarities between probabilities (and factors), the primary effect is an underestimation of the adjusted probability for much of the probability range.

Assessment

Biases can be defined as deviations from some norm (Jonathan D. Nelson, 2005; J.D. Nelson, McKenzie, Cottrell, & Sejnowski, 2010). In conjunction with producing the geospatial challenge tasks, the IARPA ICARUS program has developed metrics for assessing cognitive biases. Anchoring bias or confirmation bias in weighing evidence is assessed by a *negative entropy* metric, N and confirmation bias in seeking information is assessed using a task-specific *confirmation metric*, C .

Anchoring bias metric

Negative entropy is defined as

$$N = (H_{max} - H) / H_{max}$$

where H is the entropy of the distribution of probabilities over hypotheses and H_{max} is the maximum possible entropy. N increases with the certainty in a hypothesis (i.e., the “peakiness” of the distribution). At a given stage of updating belief probabilities [A%, B%, C%, D%] given some new INT evidence, we may assess the negative entropy, N_{ACT-R} , of the belief probabilities in ACT-R, and the negative entropy of the rational model, $N_{Rational}$. If $N_{ACT-R} > N_{Rational}$ then the ACT-R model is exhibiting a confirmation bias in weighing evidence – i.e., over-weighting evidence that confirms the most likely hypothesis. Conversely, if $N_{ACT-R} < N_{Rational}$ then the ACT-R model is exhibiting the anchoring bias.

Confirmation bias metric

Confirmation bias in seeking evidence, is assessed by the fraction, C , of SIGINT choices requested about the

insurgent group that has been assigned highest probability of being the attackers.

$$C = \frac{\text{No. of SIGINT choices on the highest Prob. group}}{\text{Total no. of SIGINT choices}}$$

SIGINT provides considerable weight when “chatter” is detected, so selection of SIGINT for the highest probability group is interpreted as being confirmatory. It is assumed that if $C > .5$ then the model exhibits confirmation bias in seeking evidence (random choice strategy by $C = .25$).

Results and Discussion

Each model was used to simulate 30,000 layer selections in 10,000 tasks. By using metrics that we explained in the previous section, we could identify that the ACT-R model exhibits anchoring and confirmation biases while conducting the task.

Anchoring bias in weighing evidence

In the geospatial task, the ACT-R model revises its probability distribution over hypotheses after each layer selection, and this can be compared against the probability distribution of the rational model. As can be seen in figure 3, the ACT-R model is most often showing lower negative entropy than the rational model ($N_{\text{ACT-R}} < N_{\text{Rational}}$). In other words, rather than showing a confirmation bias it is exhibiting a form of anchoring bias.

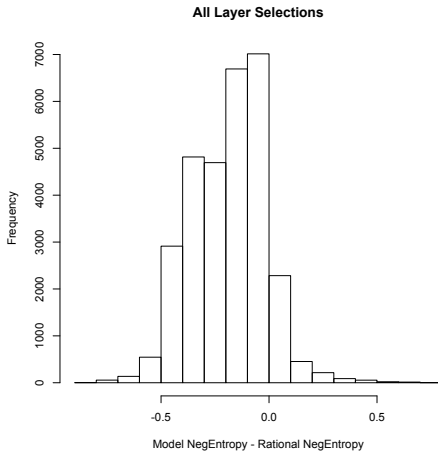


Figure 3: Difference negative entropy between the ACT-R model and rational model after each layer selections.

Confirmation bias in seeking evidence

We analyzed the fraction of SIGINT choices for which the model requests SIGINT on the group with the highest probability. The result of the fraction for the ACT-R model is presented in table 2. The fraction of the model is greater than .5, so the ACT-R model is exhibiting confirmation bias in seeking evidence according to the C metric.

We also analyzed how the INT layers selected by the ACT-R model compared to the rational choice based on the expected information gain. The result is presented in figure

4. Note that the number of alternative choices varies within a task: The task begins with seven alternatives (IMINT, MOVINT, SOCINT, and four SIGINTs) available, and depending on the selection of the layer, the alternatives decrease within each trial.

Table 2: The results of the confirmation bias in seeking evidence for both models.

	SIGINT on the highest prob. group	Total No. of SIGINT	Fraction
ACT-R Model	6,191	9,044	.68

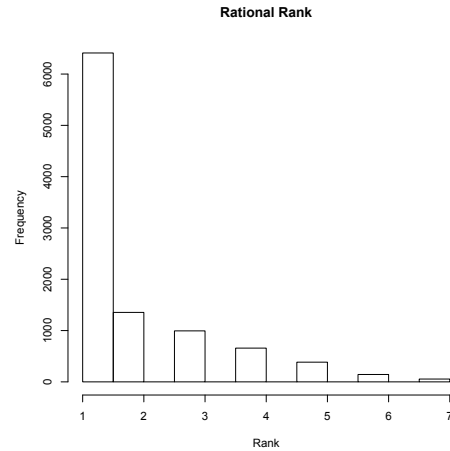


Figure 4: Frequency of the ACT-R model selecting the rational choice of Rank n (Rank 1 is the optimal choice).

Table 3 shows a confusion matrix that indicates the proportion of times the ACT-R model makes the same choice as the rational model. Although the ACT-R model agrees with the rational model at a level well above chance, it often differs from the rational. The rational model scarcely selects SOCINT layer (3 times among 30000), because the expected information gain for SOCINT is relatively low.

Table 3: Confusion matrix of the ACT-R model and rational model for layer selection.

		Rational Choice			
		IMINT	MOVINT	SIGINT	SOCINT
ACT-R Choice	IMINT	75%	15%	4%	0%
	MOVINT	18%	79%	4%	66%
	SIGINT	4%	3%	91%	0%
	SOCINT	3%	3%	1%	33%

Note that there is some interaction between the anchoring bias in evidence weighing and any biases that might emerge in choosing layers. If the ACT-R models (or participants) under-weight evidence and believe in a “less peaky”

probability distribution over hypotheses, then that can affect how far they believe that the current state or next state is from the goal, or how much more uncertainty can be reduced by a given layer choice. Biases in beliefs about the current situation will impact evidence-gathering choices.

The ACT-R model exhibits confirmation bias when evaluated against the ICArUS task-specific norm, *C*, which measures the propensity to use SIGINT to confirm the strongest current hypothesis. However, the selection of INT layers is generally highly consistent with the rational norm of seeking evidence that will produce the highest expected information gain. This illustrates how the notion of “bias” is dependent on the choice of norm, and how such norms do not always agree, especially in the case of “confirmation bias” (Jonathan D. Nelson, 2005). It has been shown (Austerweil & Griffiths, 2011) that confirmatory strategies are rational for a large class of tasks and people appear to approximate choices based on expected information gain.

Acknowledgments

This work is supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract number D10PC20021. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained hereon are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI, or the U.S. Government.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036-1060.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Austerweil, J. L., & Griffiths, T. L. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, *55*(3), 499-526.

Cheikes, B. A., Brown, M. J., Lehner, P. E., & Adelman, L. (2004). Confirmation bias in complex analyses MITRE Center for Integrated Intelligence Systems. Bedford: MA: MITRE.

Chorev, M. (1996). Surprise Attack. The Case of the Yom-Kippur War. Fort McNair: WA: The industrial College of the Armed Forces.

Convertino, G., Billman, D., Pirolli, P., Massar, J., & Shrager, J. (2008). The CACHE Study: Group Effects in Computer-Supported Collaborative Analysis. *Computer Supported Cooperative Work (CSCW)*, *17*(4), 353-393.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*, 591-635.

Grabo, C. M., & Goldman, J. (2004). *Anticipating surprise: Analysis for strategic warning*. Washington, D.C.: Joint Military Intelligence College.

Heuer Jr, R. J. (1999). Psychology of intelligence analysis: Center for the Study of Intelligence, Central Intelligence Agency (Washington, DC).

Heuer, R. J. (1999). *Psychology of Intelligence Analysis*. Washington, D.C.: Center for the Study of Intelligence.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of learning and motivation*, *32*, 385-418.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological review*, *94*(2), 211-228.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Klein, G., Moon, B., & Hoffman, R. R. (2006a). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, *21*(4), 70-73.

Klein, G., Moon, B., & Hoffman, R. R. (2006b). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, *21*(5), 88-92.

Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft*, *8*, 5-19.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Nelson, J. D. (2005). Finding Useful Questions: On Bayesian Diagnosticity, Probability, Impact, and Information Gain. *Psychological Review*, *112*(4), 979-999. doi: 10.1037/0033-295x.112.4.979

Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience Matters. *Psychological science*, *21*(7), 960-969.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175.

Pirolli, P., & Card, S. K. (2005). *The sensemaking process and leverage points for analyst technology*. Paper presented at the 2005 International Conference on Intelligence Analysis, McLean, VA.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). *The cost structure of sensemaking*. Paper presented at the INTERCHI '93 Conference on Human Factors in Computing Systems, Amsterdam.

Simon, H. A. (1974). How big is a chunk? *Science*, *183*, 482-488.

Tolcott, M. A., Marvin, F. F., & Lehner, P. E. (1989). Expert decision-making in evolving situations. *IEEE Transactions on Systems, Man and Cybernetics*, *19*(3), 606-615.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly journal of experimental psychology*, *12*(3), 129-140.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Prentice-Hall: Upper Saddle River, NJ.