Evaluating and communicating geologic reasoning with semiotics and certainty estimation

William C. Parcell[1], Lisa M. Parcell[2]

ABSTRACT

Cognitive and conceptual uncertainties are critical elements in geology from the earliest data collection stage to concluding interpretations. How a geologist conceptually weighs the importance of various data greatly influences final interpretations. In order for the process of data selection and interpretation to be transparent and repeatable, field methods and analyses should be able to communicate these cognitive processes, yet such uncertainty is difficult to characterize and estimate with standard statistical methods based on frequency probability. Semiotics and expert systems are used to frame discussion of the methods of stratigraphic reasoning and develop a field-based method to communicate cognitive and conceptual uncertainty from data collection to final interpretation. In semiotics, signs are the meanings an observer gives to an object to reach an interpretation. Uncertainty in knowledge is derived from problems in recognizing objects and judgment of inferred relevance to a larger interpretation. A final measure of geologic interpretation certainty (IC) can be derived from the product of the confidence factor (CF) of a measurement and the relevance factor (RF) of that object to an interpretation. Communicating levels of interpretation certainty, relevance, and significance allows geologic investigations to be more transparent to subsequent geoscientists, non-geologists, and even the original investigator.

INTRODUCTION

The ability of geologists to derive accurate conclusions from stratigraphic data is limited by the presence of systematic and conceptual uncertainty. Uncertainty can be classified as "reducible" and "irreducible." Natural uncertainty is "inherent" or irreducible, whereas data and model uncertainty contain both reducible and irreducible components. Irreducible uncertainty in data and models is largely a result of the presence of natural uncertainty, including randomness and chaos. Reducible uncertainty is a result of a lack of knowledge of a system. Reducible uncertainty (Fig. 1) emerges at the empirical level, (e.g. measurement or sampling error), the cognitive level, (e.g. vagueness and ambiguity in natural language/signs), and the conceptual level (e.g., uncertainty regarding relationships between data).

Reducible uncertainty at the cognitive and conceptual levels in sedimentology and stratigraphy has received some attention in recent years (Dott, 1998; Baker, 1999, 2000; Frodeman, 1995; Raab and Frodeman, 2002) and is the focus of this paper. Stratigraphy and sedimentology (and geology in general) run counter to the traditional view of the method of science that is often portrayed as essentially experimental in nature. Dott (1998), Baker (1999, 2000) and Frodeman (1995), among others, posit that geology is one of the few modern sciences that maintain a holistic method of inquiry that stresses "synthetic reasoning" to interpret Earth processes rather than experimental design. Synthetic reasoning is a sequence of mental steps that individuals follow in assembling components into an effective, coherent system (Lee and Johnson-Laird, 2005). While the experimental sciences endeavor to understand fundamental principles and from this knowledge make predictions, geology along with cosmology, archaeology, and evolutionary biology, seek not to predict what will happen in the future, but instead anticipate what will be discovered about the past (Primack and Abrams, 2006). Much of the geologic method involves reasoning back from the existence of clues towards a hypothesis to explain their presence and relationships (Frodeman 1995). These clues (or signs) include rocks, sediments, fossils, and other suggestions of Earth processes. Much of uncertainty in stratigraphy is derived from whether or not the geologist recognizes the signs that lead to an interpretation of causation.

In traditional field geology, cognitive and conceptual uncertainties become critical elements at the earliest data collection stage. A fundamental component of fieldwork is the selection of appropriate data. Which criteria to use and how a geologist weighs and selects the importance of various data critically influence final interpretations and conclusions. What criteria are used to select data depends on the context of the situation. Selection of data includes both perception of the data and placing it within a conceptual model; in short, the geologist is already interpreting at the moment of data collection. But the manner in which a geologist traditionally approaches data collection can be seen as running counter to established accounts of the scientific method. As discussed by Rabb and Frodeman (2002), the standard description of the scientific method is largely based on laboratory sciences (e.g. physics and chemistry) and presumes that measurements of an experiment must be indirect (or disembodied), positing that perceptions of the observer should have no affect on the outcome of the result. However, in traditional geologic investigations, much of the data collection process directly invokes geologists' perceptions of the data. Hence, geology has been often described in terms similar to detective work, where the geologist must interpret the scene of a crime. Small bits of evidence remain to construct a large, encompassing story that explains how the pieces ended up in their particular geometry. To solve the crime, geologists must use various tools at their disposal (Frodeman 2000, Raab and Frodeman 2002). These include selecting what is deemed important information based on professional training,

[1]Department of Geology, Wichita State University, 1845 Fairmount Ave., Box 27, Wichita, KS 67260; william.parcell@wichita.edu
[2]Elliott School of Communication, 1845 Fairmount Ave., Box 31, Wichita State University, Wichita, KS 67260; lisa.parcell@wichita.edu

conceptual models (such as use of analogs, uniformitarianism, catastrophism, sequence stratigraphy, cyclostratigraphy, etc), logic rules, and judgment.

In order for the process of data selection and interpretation to be truly transparent and repeatable, the method must be able to communicate these cognitive processes. This paper presents a method to use semiotics and application of degrees of significance to calculate and communicate the degree of confidence that a geologist places on the selection and interpretation of data. The challenge is to derive a relatively simple, rapid, intuitive method to apply in a classroom or professional setting, but also based on sound philosophic and semiotic theory. The theory and method described in this paper focus on sedimentology and stratigraphy, but could be just as well applied to other field disciplines in geology.

PERCEPTION AND CONCEPTION

While geology is primarily concerned with the results of scientific studies, it is necessary that the discipline recognize the importance and influence of geologists' reasoning processes. Simply by observing an object of study, the geologist influences the nature and meaning of the object. The field geologist is both the data-collecting instrument and interpreter of that data. As such, s/he is not an independent observer.

Interpretation begins at the data collection stage. A geologist approaches an outcrop with the intention of collecting information to observe, derive, or test a hypothesis or fit the data within a holistic model. The geologist also comes to the outcrop with cognitive approaches influenced by training, conceptual models, and accepted logic rules. The ensuing cognitive processes of interpretation are typically not portrayed in scientific results because communication of this information is either too difficult to represent, not recognized by the geologic community, or there is an aversion to conceding that these judgments affect observations and results.

For many years, there have been declarations against the overuse of conceptual models in geology. Ager (1970) for instance argues that many geologists travel the world merely to confirm their own ideas and simply ignore everything else. Such arguments warn against adding an interpretive overlay to empirical classification and correlation. Miall (2004) and Miall and Miall (2004) state that geologists must maintain a grounded perspective in "descriptive facts" and not be driven by conceptual models. The crux of such arguments is that traditional field-based geology is descriptive in nature and essentially detached from the geologist's influence. This essentially posits that interpretation should ideally enter the picture only after description and sampling of the rock. Many scientists may also remain reluctant to accept that human factors influence scientific discovery. However, many philosophers of science, neuroscientists and semioticians would argue that as humans we cannot escape the reality that we are influenced by our background and cognitive processes (Lakoff and Johnson, 1999; Burton, 2008). Interpretation begins the moment a geologist selects an area of study or selects a sample at the outcrop. As Raab and Frodeman (2002) recognize, a field geologist is must subjectively decide what specimens meet their standards for study. Interpretation is present throughout the geologic rendering of the scientific method. They suggest that two interacting cognitive processes of perception and conception are functioning when geologists collect field data. Perception is the process of recognizing characteristics (or signs) in objects that indicate larger meaning. Conception is the act of placing these clues together forming or expanding a conceptual framework.

Through the perception and conception phases, a number of factors influence the accumulation of knowledge and can be viewed as a progressive hierarchy of uncertainty. Perception of objects is influenced by observer's prior knowledge of subject, the presumed goal of project, the set of skills at one's disposal, the accuracy of measurements, and the classification system used (including the technical language used to describe the object). Not all of these can be rationally established. However, it is possible to estimate measurement accuracy and confidence in classification in the field. Weighting the confidence in the measurement is a gauge of how well the observation fits within the observer's set of tools and preconceptions. These processes can also be viewed

HIERARCHY OF UNCERTAINTY

LEVELS OF UNCERTAINTY				
High Level	Irreducible Uncertainty	Reducible Uncertainty		
		Data/Parameter Uncertainty	Model Uncertainty	
Mid Level	Natural Uncertainty	Empirical Uncertainty	Cognitive Uncertainty	Conceptual Uncertainty
Low Level	Randomness Chaos	Randomness Inaccuracy Imprecision Vagueness	Ambiguity Undecidedness Vagueness	Ambiguity Undecidedness Vagueness

FIGURE 1. Heirarchy of uncertainty level and types. The highest level of uncertainty is a division of reducible and irreducible (natural) uncertainties. Reducible uncertainties can be viewed as occurring at the empirical, cognitive and conceptual stages of an investigation, each characterized by different types of uncertainty including randomness, inaccuracy, imprecision, vagueness, ambiguity and undecidedness. For a thorough discussion of types of uncertainty, see Klir and Wierman (1999) and Berkan, and Trubach, (1997).

within the framework of hermeneutics and semiotics - the study of signs (Eco 1976, Eco and Sebeok 1988).

SEMIOTICS AND EXPERT SYSTEMS AS FRAMEWORK FOR STRATIGRAPHIC REASONING

Because the geologist's human nature influences the meaning of the objects under examination, the manner in which an individual selects data is critically influential on the interpretation of that data. However, recognizing this potential source of uncertainty is just the first step. Ideally, if one wants to make the study more replicable, one must also be able to communicate the ensuing cognitive processes - with all resulting uncertainty. If one cannot portray when and what data were deemed more important than others, then much of the interpretation is missing. The challenge is deriving a relatively rapid and intuitive method to portray these judgments in the field. The communications theory of semiotics along with the artificial intelligence techniques of expert systems provide useful tools to frame discussion of a geologist's reasoning and, when combined, provide a usable system to communicate uncertainty levels in judgments.

Semiotics and hermeneutic processes

Hermeneutics is the study of interpretation, including how and why an observer places meaning to objects and ideas. Proponents of modern hermeneutics posit that an observer's understanding of an object encompasses many aspects of his or her background experience and training. Within modern hermeneutics, semiotics is the study of signs and how meaning arises by the processes of perceiving and conceiving (i.e. interpreting) objects.

In semiotics, "signs" can be defined as something present that represents something absent (Leeds-Hurwitz 1993). In semiotics, a sign is a set of characteristics that represents an object to somebody in some capacity. Characteristics of an object govern what sign or signs are recognized. In turn, a sign creates in the mind of a person an equivalent or more developed sign or idea (Fig. 2A), the "interpretant" (Chandler 2007). For a simple example, a rock (i.e. an object) may be characterized by certain properties (or signs) that have meaning to a geologist and can lead to an interpretation. An object is an identified by commonly understood, shared properties (the Realm of Common Reality in Fig. 2B). In our example, the object is commonly understood by the general public to be a rock. The rock characteristics then have additional meaning or relevance to a geologist. This rock may have particular characteristics (signs) such as texture, mineralogy, color, etc., recognized by geologists that point them to an interpretation or greater meaning. For our particular example, the geologist may perceive various signs including medium gray color, spherical grains with concentric layers, mud-sized grains between the spherical grains, calcium carbonate (from an HCl test), and a packing pattern that suggests a grain-supported fabric. To reach an interpretation from these signs, geologists place the rock's particular characteristics within the framework of their education, prior experience, conceptual models, hypotheses, etc. A geologist's final interpretation of the rock's (object's) signs may be to classify it as an "oolitic packstone" (Realm of Meaning and Implication in Fig. 2B). In semiotic terms, the implication of the characteristics (signs) of an object is called the "interpretant." It is important to note that the sign occurs as the intersection between the realm of common reality and realm of meaning and implication. This concept is illustrated in Figure 2B by the overlapping circles of both realms. The influence of both 1) shared, common reality and 2) hypothesis and conceptual models influences what signs are recognized. From recognition of rock types to interpretations of depositional environment, paleoclimate and cyclicity, the cognitive process of object -> sign -> interpretant can be applied throughout geologic investigations. Figure 3 illustrates a simple example of signs and groups of signs pointing a geologist towards higher and higher orders of interpretation.

Through each object to sign and sign to interpretant step, uncertainty in measurement, perception, judgment, and interpretation influences subsequent analyses. However, to date, there is no technique to recognize and communicate these cognitive uncertainties into the interpretation process.

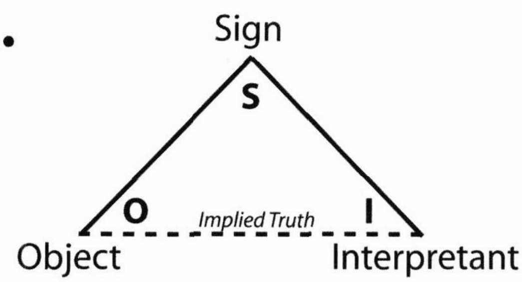
Expert systems

Expert systems can synthesize the elements of philosophy, psychology, and communications within a particular science. An expert system contains knowledge derived from an expert in some narrow field. The primary application of expert systems research has been to make expertise available to decision makers and technicians (Nikolopoulos, 1997).. These knowledge-based applications of artificial intelligence have enhanced productivity in business, science, engineering, and the military (Luger, 2002). However, advances in expert systems also place a mirror to the experts themselves, obliging the expert to explain how they reason and arrive at conclusions.

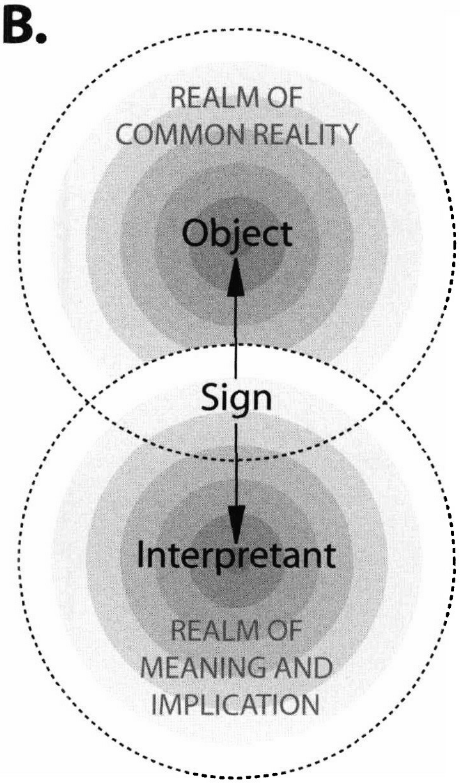
An expert is one who possesses specialized skill, experience, and knowledge (in this case a geologist) along with the ability to apply general or specialized logic rules to efficiently resolve a problem (Harmon and King, 1985). An expert system often includes tools that aid the design, development, and testing of the knowledge base, just as a geologist comes to an outcrop with certain skills at their disposal. Expert systems consist of three elements - a knowledge base, a user interface, and a reasoning engine (typically IF- THEN logic rules). The knowledge base contains "control information" for an expert system, which may include pre-defined concepts, classifications, and logic rules.. This is the equivalent of a geologist's learned knowledge and professional training. The user interface of the expert system is traditionally used to pose questions to the non-expert about potential problems for the system to then resolve. Questions might include "what is the likely geologic setting that produced the given suite of rocks," or "where is the best location to drill for petroleum or water resources."

This impartial, detached interface can also be turned around and used to ask the expert questions about their

A.



B.



SEMIOTIC RELATIONSHIPS	COGNITIVE PROCESS	TYPES OF UNCERTAINTY
Object: Item identified by understood common, shared characterization or classification	Identify	Randomness Inaccuracy Imprecision Vagueness
Sign: Meaning of object within particular context; context based on education, prior experience, conceptual model, hypothesis, theory, etc.	Perceive	Ambiguity Undecidedness Vagueness
Interpretant: Implications of object meaning; how knowledge of object expands larger understanding	Conceive	Ambiguity Undecidedness Vagueness

FIGURE 2. Semiotic relationships between object, sign and interpretant. Upper diagram (A) displays the classic triangle illustration of the continuous association between object to interpretant by Ogden and Richards (1923) and explained by Eco (1979). Signs can be viewed as characteristics of an object suggesting a particular meaning (the interpretant). Lower diagram (B) represents contextual influences on semiotic relationships. The semiotic element "sign" occurs at the intersection of the realm of common reality and realm of meaning and implication. This coincides with the cognitive processes of perception and conception. Perception of an object occurs in the realm of common reality by recognition of commonly understood properties or characteristics. Conception of that sign relates to the projection of meaning to that object. Various types of uncertainty are associated with identifying objects (randomness, inaccuracy, imprecision, vagueness), signs and interpretants (ambiguity, undecidedness, vagueness). Irreducible uncertainty is not considered in this paper.

confidence in certain observations or interpretations. The reasoning engine then implements IF-THEN logic rules and controls the interview process. The logic rules can be run by any number of possible operations including Boolean, propositional, multi-valued (fuzzy), etc. The following section discusses the integration of uncertainty factors into a semiotic and expert systems framework.

Confidence factor (CF)

The education, prior experiences, conceptual models, hypotheses, and personality that geologists bring to the field all affect their ability to recognize signs. Therefore, two experts at an outcrop may not immediately recognize (or "perceive") the same signs. This is related to two factors: 1) the confidence of the observer in their observation or varying classifications for a particular object (as influenced by field conditions and training/skill level), modified by 2) the observer's personality. Expert

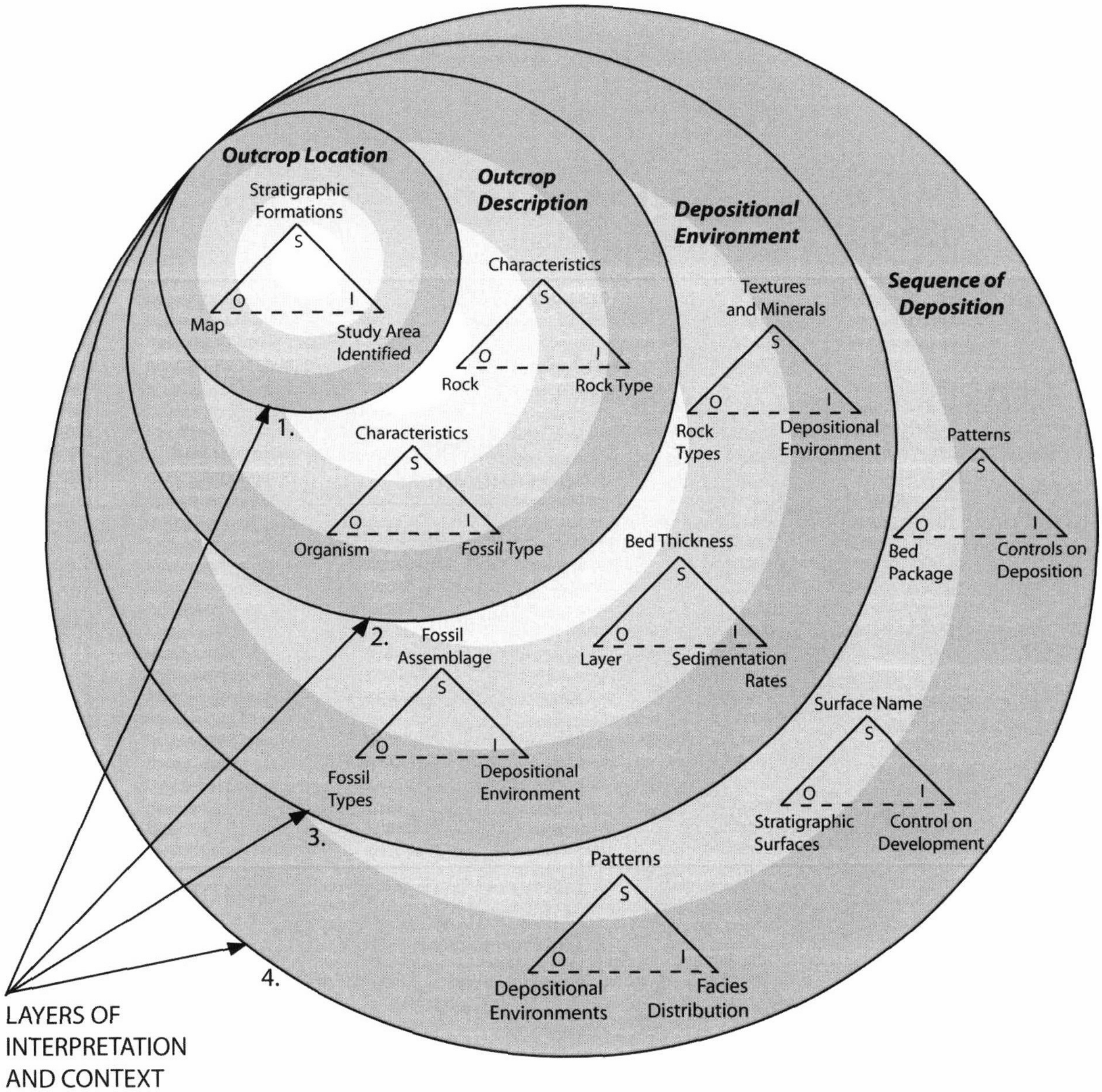
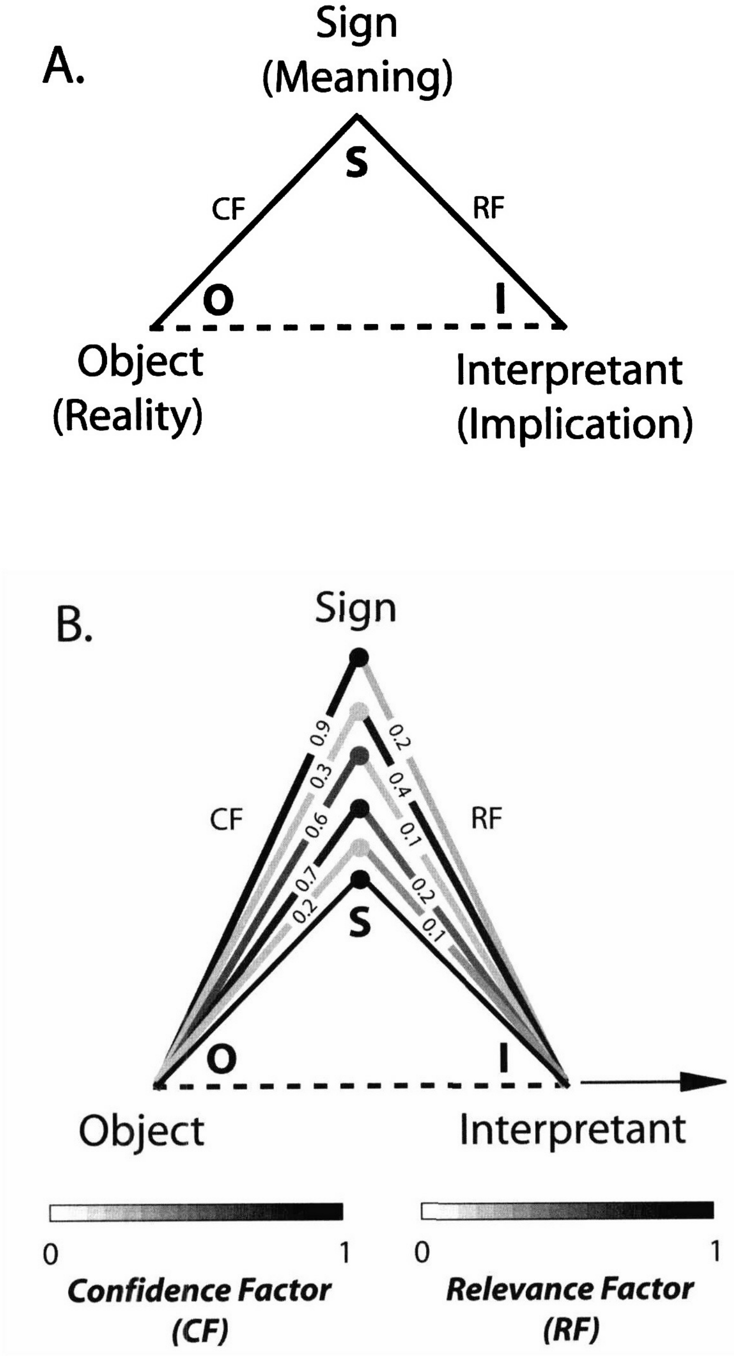


FIGURE 3. Continuous semiotic connection in a stratigraphic study from data collection (upper left) through final interpretation (lower right). With each interpretation, the potential exists to increase background context as well as levels of uncertainty. While the diagram begins with the process of map interpretation, this is only for illustration purposes. The actual heurmeneutic process of semiotics allows the perception of signs anywhere in this limited diagram and perception in any order depending on the observer's background understanding.

systems are designed to deal with uncertain and vague linguistic information. Expert systems handle these problems in different ways, often by permitting users to represent their degree of confidence in a measurement with a numerical scale. Use of degree of confidence and degree of relevance are both used in expert systems. The confidence factor (CF) is a measure of the certainty assigned to a measurement or description. Certainty factors are often expressed as a value (0 to 100%) or (0 to 1.0) where 100% or 1.0 implies that the attribute's value is known with absolute certainty. These numbers are similar in nature to probabilities, but instead of following the mathematical definitions used in calculating probabilities, these values are meant to communicate the levels of confidence in human reasoning. Within semiotics, the confidence factor may be viewed as a measure of uncertainty affecting what signs are recognized from objects (Fig. 4). Each CF is independent from other CF's that affect an observer's recognition of particular signs. It must be emphasized that estimation of CF is not a value derived from an impartial witness or instrument, but a quantity inclined by the observer's honest reflection of their own scientific observations. Some may argue that a non-metric estimation of certainty is a highly subjective judgment, insinuating that it is a meaningless, biased enterprise. However, "objectivity", as Gould (2000, p.104-105) states, "cannot be equated with mental blankness; rather, objectivity resides in recognizing your preferences and then subjecting them to especially harsh scrutiny." As many philosophy, neuroscience, and psychology studies

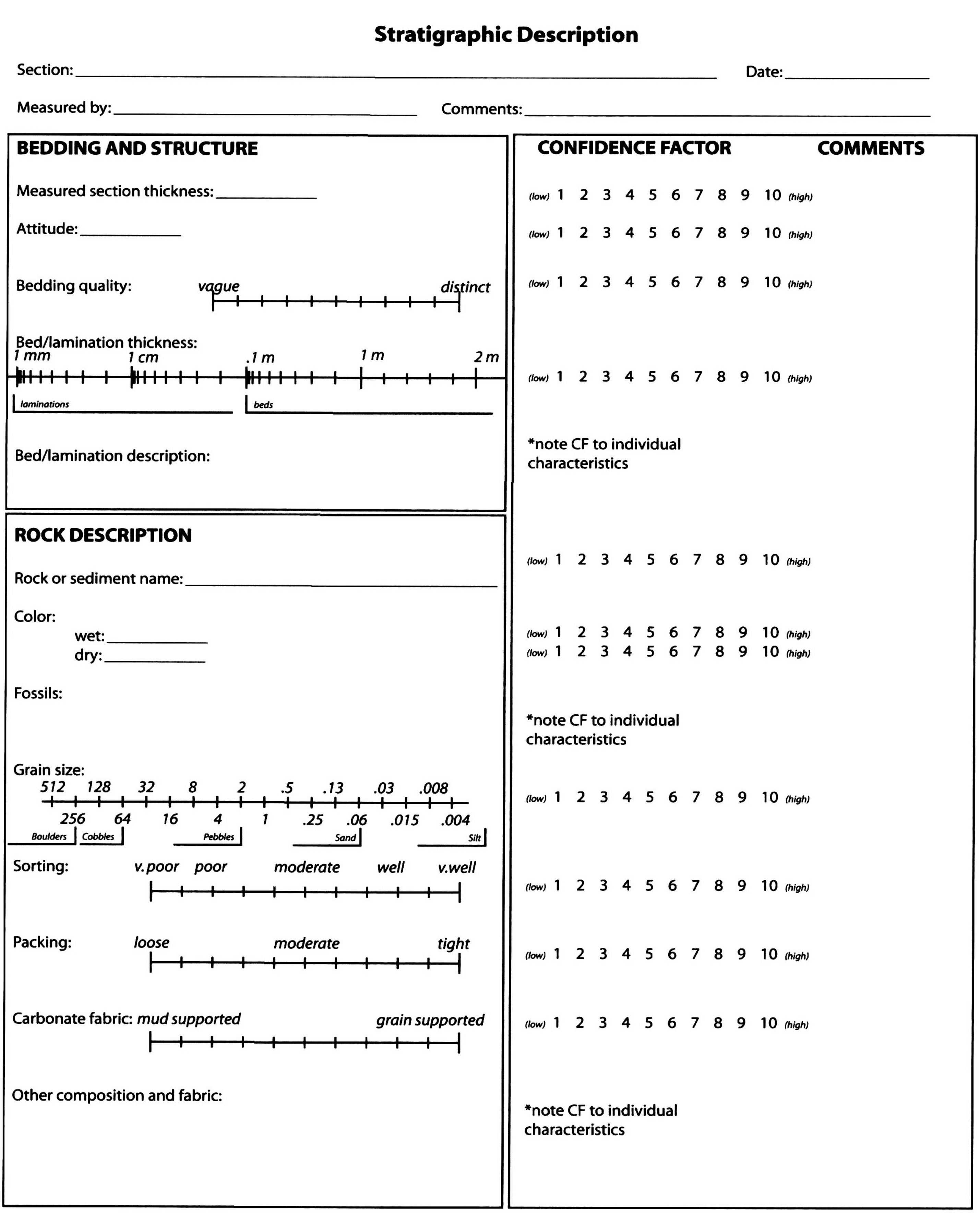


CF * RF	= IC
0.9 * 0.2	= 0.18
0.3 * 0.4	= 0.12
0.6 * 0.1	= 0.06
0.7 * 0.2	= 0.14
0.2 * 0.1	= 0.02
	IC_S = 0.52

FIGURE 4. The position of the confidence factor (CF) and the relevance factor (RF) within the semiotic scheme (A). Interpretation certainty (IC) is derived from the product of the CF of each object and the RF of each sign (B). The variable height of the sign vertex does not imply any change in the CF, RF or IC. Change in height is for clarification purposes only.

of human behavior have shown (e.g. Kuhn, 1962; Burton, 2009; Lakoff and Johnson, 1999), attempts to be "scientifically" objective are dramatically influenced by human cognition. Human physiology does not permit the idealized "objective" thought. Cognitive processes arise out of involuntary brain mechanisms that operate independently of reason (Burton, 2009). Recent research in neuroscience has shown that estimates of confidence in knowledge are biologically controlled (through genetic and environmental factors) and largely separate from rational thought (Bouchard et al., 1990; Burton, 2009). Yet, these "feelings" of confidence combined with rational thought influence the decision making process. These estimations of confidence are usually not acknowledged in those sciences that rely on direct human observation of data (e.g. archeology, geology, ecology, etc.). Geologists perform these self-analyses (consciously or subconsciously) during everyday observations and interpretations of data, but these tacit assumptions often go unstated. Semiotics, hermeneutics, and expert systems provide a means to communicate these internal judgments.

There are potential pitfalls with any system that estimates confidence or uncertainty based on self-analysis. Users' personalities may influence the designation of CF's. Confident individuals may tend to consistently specify high CF's and more meek personalities tend towards lower CF's. However, by communicating these confidence rankings, readers may be able to recognize and



Stratigraphic Description

Section: ______________________ Date: ____________

Measured by: ______________________ Comments: ______________________

BEDDING AND STRUCTURE

Measured section thickness: __________ (low) 1 2 3 4 5 6 7 8 9 10 (high)

Attitude: __________ (low) 1 2 3 4 5 6 7 8 9 10 (high)

Bedding quality: *vague* — *distinct* (low) 1 2 3 4 5 6 7 8 9 10 (high)

Bed/lamination thickness: *1 mm* *1 cm* *.1 m* *1 m* *2 m* (low) 1 2 3 4 5 6 7 8 9 10 (high)

laminations *beds*

Bed/lamination description:

*note CF to individual characteristics

ROCK DESCRIPTION

Rock or sediment name: ______________________ (low) 1 2 3 4 5 6 7 8 9 10 (high)

Color:

wet: __________ (low) 1 2 3 4 5 6 7 8 9 10 (high)

dry: __________ (low) 1 2 3 4 5 6 7 8 9 10 (high)

Fossils:

*note CF to individual characteristics

Grain size: *512 128 32 8 2 .5 .13 .03 .008* / *256 64 16 4 1 .25 .06 .015 .004* (low) 1 2 3 4 5 6 7 8 9 10 (high)

Boulders | Cobbles | Pebbles | Sand | Silt

Sorting: *v.poor poor moderate well v.well* (low) 1 2 3 4 5 6 7 8 9 10 (high)

Packing: *loose moderate tight* (low) 1 2 3 4 5 6 7 8 9 10 (high)

Carbonate fabric: *mud supported* — *grain supported* (low) 1 2 3 4 5 6 7 8 9 10 (high)

Other composition and fabric:

*note CF to individual characteristics

CONFIDENCE FACTOR — COMMENTS

FIGURE 5. Example of outcrop check sheet for describing measured sections with confidence factor (CF) scale available for most instances. Similar results could be achieved by noting CF within field notes. Modified from Compton (1985).

acknowledge the thought process and human influence behind data collection.

Figure 6 illustrates a possible field notation system for estimating confidence factors (CF). For each observation the geoscientists' confidence in their observation for a particular object are noted on a scale from 0 to 1 (or 0 to 100%). For example, Geologist #1 might include confidence factors in field analyses by describing an outcrop feature as a "gray (1.0 CF), thin bedded (0.8 CF), fossiliferous (0.7 CF), CaCO3 (1.0 CF), and densely packed (0.7 CF)." Geologist #2 approaching the same feature might describe it as a "gray-tan (0.9 CF), thin bedded (0.9 CF), peloidal (0.6 CF), fossiliferous (0.8 CF), CaCO3 (1.0 CF), and moderately packed (0.7 CF)."

Relevance factor (RF)

Even when given the same recognized datasets, two experts may arrive at different interpretations (or "conceptions" using terminology from Raab and Frodeman, 2002). One observer may emphasize particular signs while another may recognize different characteristics as critical. An interpretation may be as simple as identification of a rock (based on fabric, grain size, mineralogy, etc) or more complex such as recognizing characteristics of stratigraphic surfaces and units (sequence boundaries or cycles) and drawing correlations between stratigraphic sections. In order to approach a more repeatable interpretation, expert systems provide a mechanism for portraying a geologist's reasoning process by asking the expert to assign a degree of relevance, or relevance factor (RF), to each observation as it pertains to a particular interpretation. Within a semiotic point of view, the relevance factor can be regarded as a measure of the inferred importance a particular sign has on an interpretation (Fig. 4). Again, this is the judgment of the observer; but, indirectly, assignment of a relevance factor is a measure of the influence of prior knowledge of subject, goal of the project, and skill set on the observers reasoning process. Akin to CF's, relevance factors can be expressed as a value (0 to 100% or 0 to 1.0) where 100% or 1.0 implies total relevance to an interpretation. However, unlike CF's, because each relevance factor relates to the same interpretant, the sum of RF's must add up to 100% (or 1.0). This can be achieved by normalizing RF's in post-fieldwork processing after the interpreter makes relevance judgments.

Returning to the field example above, the two geologists that collected the outcrop features may apply different relevance factors to their observations when making a particular interpretation. Geologist #1 might assign the name "fossiliferous grainstone" to the data collected above. She assigns the following relevance factors to the data (signs) collected as it applies to the interpretation: gray (0.1 RF), thin bedded (0.0 RF), fossiliferous (0.3 RF), CaCO3 (0.3 RF), densely packed (0.3 RF). Geologist 2 might assign the name "peloidal and fossiliferous packstone" to his observations. He assigns the following relevance factors to his data (signs): gray-tan (0.0 RF), thin bedded (0.0 RF), peloidal (0.2 RF), fossiliferous (0.3 RF), CaCO3 (0.3 RF), moderately packed (0.2 RF). Again, because each relevance factor relates to the same interpretant, the sum of RF's must equal 1.

Constructing a workable field system

Figures 5 and 6 illustrate examples of potential applications of CF and RF notations within field descriptions. By first noting confidence in object measurement and later its conceptual relevance, the geologist can communicate and retain more of the cognitive process of data collection to achieve greater transparency of method and to support subsequent field checking.

Figure 5 illustrates an example of field notes and check sheets. The use of confidence factors is dominant at the data collection phase. As discussed above, geologists recognize signs from the wide variety of objects in front of them. This requires noting and communicating the CF of measurement and classification of these objects. As such, for each description, a CF is noted next to the measurement, along with comments as to why this factor was chosen. Figure 6 illustrates both a typical stratigraphic column, which synthesizes the data collected during a particular outing, and subsequent stratigraphic interpretations of the column. Here both confidence factors and relevance factors are used to guide and communicate the process of interpretation. While data may or may not be measured to high accuracy (CF), those data that support a favored interpretation should be given high RF's, while data that are problematic for a favored interpretation should be given low RF's. Data that are confusing or poorly understood in the context of a given interpretation would rank low in terms of RF.

Post-fieldwork interpretation processing

At the conclusion of fieldwork, the geologist begins to reflect upon the data collected. This may be when s/he returns to basecamp at the end of the day, to the office at the end of field season or years later. This private reflection often includes an unacknowledged estimation of certainty in observations and the relevance of data to certain hypotheses and models. This process leads to an estimation of confidence in interpretations, however it is never reported or communicated.

This private, post-field work processing can be emulated by the processing of the objects' CF's and RF's through mathematical operations that produce a value that communicates certainty in interpretation. Figure 6 displays this processing and confidence in interpretations. The first variable in the calculation of certainty in an interpretation is the confidence factors (CF) of signs recognized from objects. The second variable in certainty processing is the relevance factor (RF) of signs that lead to an interpretation. The relevance of an object to an interpretation is only as significant as the observer's confidence in their recognition of that object. Put another way, an observer's confidence in their final interpretation can only be as strong as the component object RF's as modified by the CF's of these objects. These cognitive processes can be emulated with post-field "uncertainty processing." These self-determined values are a reflection of the cognitive processes that affect interpretation of

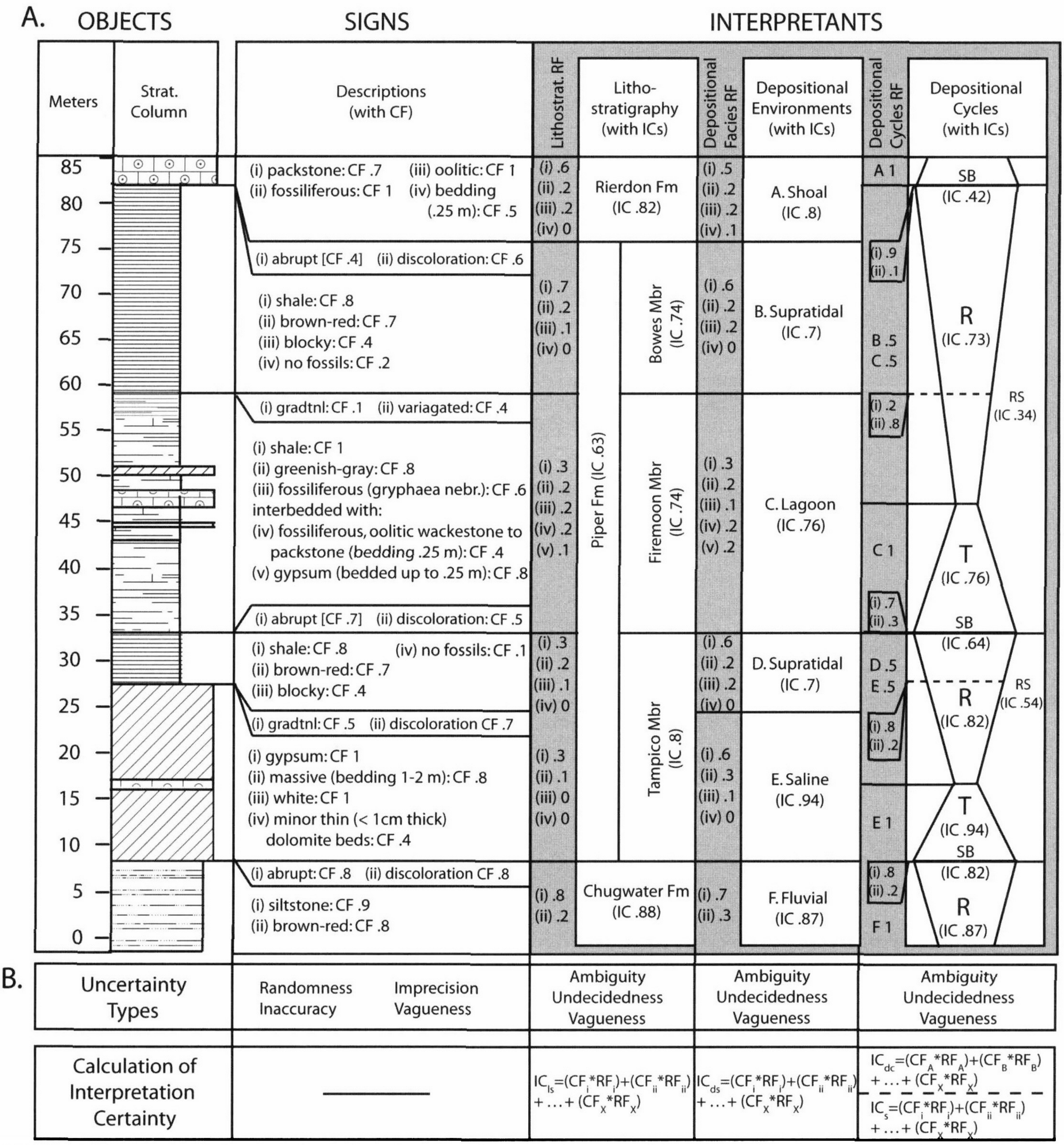


FIGURE 6. Example of the use of certainty factors, relevancy factors and interpretation certainty within a graphic stratigraphic log and associated descriptions and interpretations. See text for explanation. Stratigraphic log and descriptions derived from Parcell and Williams (2005).

physical reality. Uncertainty processing is the modification of relevance factors by confidence factors to attain an estimate of confidence in interpretation. This is achieved by multiplying a sign's relevance factor by the confidence factors of its object (Fig. 4). The result is an adjustment of the relevance factor by its associated measurement uncertainty, or simply

$$CF_X * RF_X = IC_X \qquad \text{Eq.1}$$

where IC_X is the interpretation certainty for a given object

-> sign -> interpretant analysis. The value of the IC will remain as high as its associated RF only if the CF is 1.0. Otherwise, the value of the IC will drop in proportion to the CF.

Typically, most interpretations will include multiple CF's and RF's. Therefore, an interpretation will likely result in multiple IC's. The sum of these IC's will give the interpretation certainty, or IC_S (Fig. 5).

$$IC_1 + IC_2 + \ldots + IC_X = IC_S \qquad \text{Eq.2}$$

The IC_S is the overall confidence the geologist has in an interpretation and it and its CF and RF components describe the contribution of observations to an interpretation. Data that support a favored interpretation will appear with both high CF and RF associated with a high IC_S. Data that are problematic to a favored interpretation would maintain a high CF but a be associated with a low RF and low IC_S. Data that are confusing to the observer (which may or may not fit into a favored interpretation) will be indicated with a low CF, low RF, and low IC_S. These values can then be used to communicate uncertainty in geologists' interpretations as represented in Figure 6.

Figure 6 shows an example of certainty factors, relevancy factors, and interpretation certainty displayed within a graphic stratigraphic log and associated descriptions and interpretations. Certainty factors (CF) are noted for each element of a description. For interpretation of the data, a relevance factor (RF) is recorded for each element of a description. Calculation of the interpretation certainty (IC) is achieved by summing the products of the CF and RF for each description.

Figure 7 incorporates confidence attributes (IC) onto a typical stratigraphic cross section. Here interpretation certainty of lithologies is represented by varying the degree of transparency of lithologic symbols. This provides a rapid, intuitive means to communicate the author's level of understanding.

CONCLUSIONS

For scientists, knowledge is derived from observation of physical phenomena through human cognitive processes. Attainment of knowledge is influenced by confidence in truths (i.e. interpretation) and methods of approaching truth (including empiricism, rationalism, and constructivism). The system described here allows communication of these cognitive processes to other researchers in order to be more transparent and useful. Geologists are connected with their subject matter at the cognitive level. In this historical science where interpretation of objects is required at the earliest stages of data collection, it is impossible to measure and describe an object without influencing the selection and meaning of an object through past experiences, working hypotheses, personality, and professional training.

The application of semiotics within an expert systems context provides a framework for examining levels of user confidence and uncertainty. By communicating levels of certainty, relevance, and significance, the process of

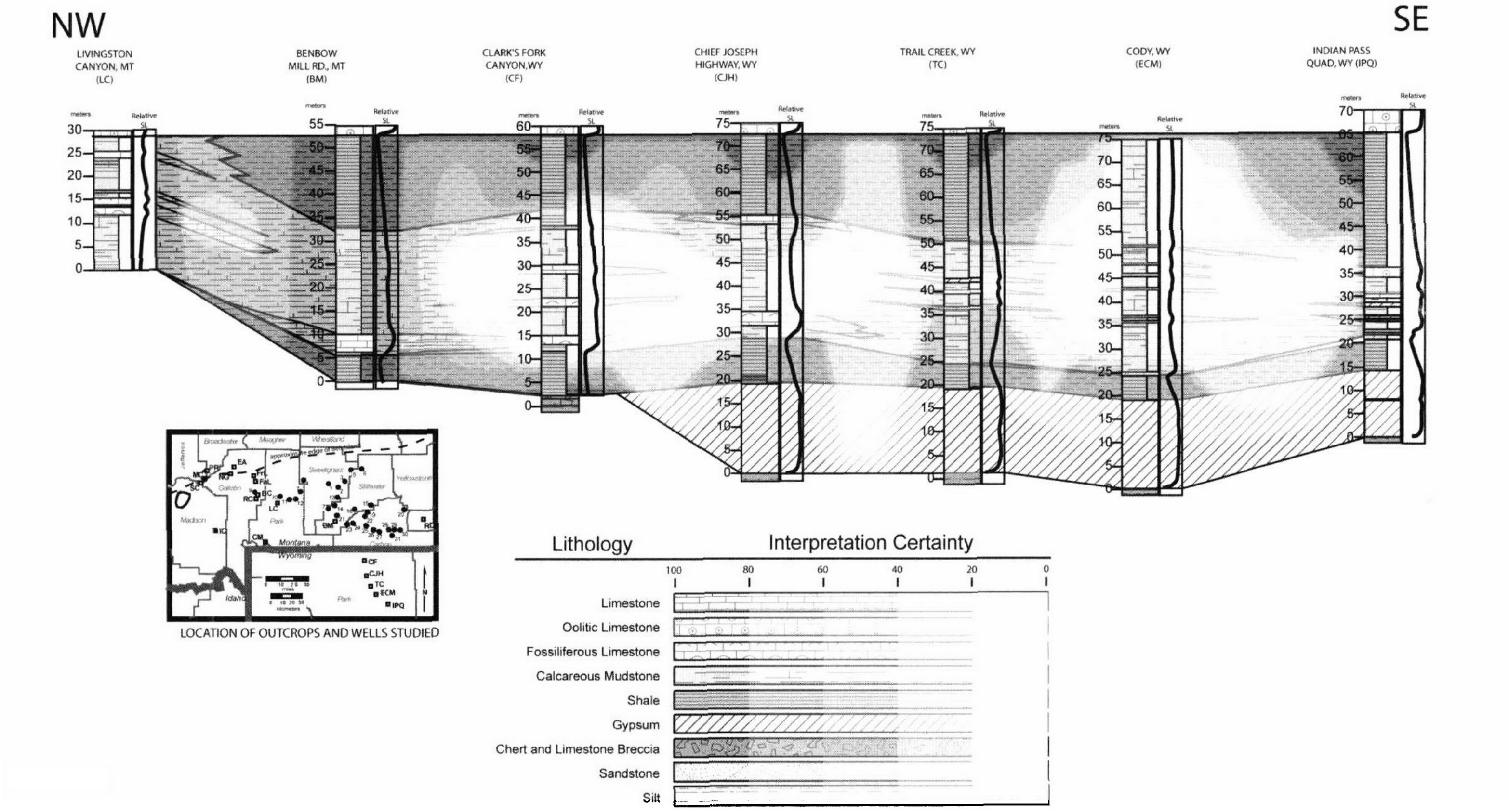


FIGURE 7. Overlay of confidence attributes (IC) onto a stratigraphic cross section. Interpretation certainty of lithologies is represented by varying the degree of transparency of lithologic symbols.

recognizing *information* is *more* transparent to subsequent geologists, non-geologists (e.g. engineers and project team members), and even to the original investigator. For the geologist, these measures provide valuable insight into their colleagues' research. Understanding a researcher's confidence in the observations and interpretations for a particular rock strata, in addition to how and where data were gathered for that strata, gives subsequent researchers a more solid understanding of the work done in the area. An area of future study would be a comparison or analysis of multiple geologists ICs on a single strata, perhaps leading to a more accurate IC for the strata.

For the non-geologist, such as engineers and project team managers, this measurement of confidence provides a tangible estimate useful in predicting risk. Engineers and product mangers for an oil exploration unit, for example, may not fully understand or recognize the importance of the various signs (objects) the geologist observes in the field, yet they rely on petroleum geologists' final interpretations to decide when and where to look for natural resources. A geologist's interpretation with an IC of 0.45 might raise a red flag, for example, while an IC of 0.95 would be considered much less risky. Future studies in this area could examine the impact of IC on non-geologists in the areas of oil and mineral exploration, environmental assessment, public policy, and climate change.

REFERENCES

Ager, D.V.,1970, On seeing the most rocks: Proceedings of the Geologists' Association, v. 81., p. 421.-427.

Baker, V.R., 1999, Geosemiosis: Geological Society of America Bulletin, v.111, no. 5, p. 633-645.

Baker, V.R., 2000, Conversing with the earth: the geological approach to understanding, *in* Frodeman, R. ed., Earth Matters: The Earth Sciences, Philosophy, and Claims of Community, Prentice Hall, 209 p.

Berkan, R.C., and Trubach, S.L., 1997, Fuzzy Systems Design Principles: Building Fuzzy IF-THEN Rule Bases: Piscataway, N.J., IEEE Press, 496 p.

Bouchard, T.J., Lykken, D.T., McGue, M., Segal, N.L., Tellegen, A., 1990, Sources of human psychological differences: the Minnesota Study of twins reared apart: Science, v. 250, n. 4978, p. 223-228.

Burton, R., 2009, On Being Certain: Believing You Are Right Even When You're Not: St. Martin's Press, 272 p.

Chandler, D., 2007, Semiotics: The Basics: Routledge Press, 328 p.

Compton, R.R., 1985, Geology in the Field: John Wiley & Sons, 398 p.

Dott, R.H., Jr., 1998, What is unique about geological reasoning?: GSA Today, v.10: p.15-18.

Eco, U., 1979, A Theory of Semiotics: Bloomington, Indiana University Press, 354 p.

Eco, U. and Sebeok, T., 1988, The sign of three: Dupin, Holmes, Pierce: Bloomington, Indiana University Press, 236 p.

Frodeman, R., 1995, Geological reasoning: Geology as an interpretive and historical science: Geological Society of America Bulletin, v. 107, no. 8, p. 960-968.

Frodeman, R., 2000, Earth Matters: Philosophy and Geology: Prentice Hall, 240 p.

Gould, S. J., The Lying Stones of Marrakech: Penultimate Reflections in Natural History: Harmony Books, New York, 384 p.

Harmon, P. and King, D., 1985, Expert Systems: Artificial Intelligence in Business: John Wiley & Sons, 283 p.

Kuhn, T., 1996, The Structure of Scientific Revolutions: University of Chicago Press, 226 p.

Klir, G.J. and Wierman, M.J., 1999, Uncertainty-Based Information: Physica-Verlag Heidelberg (2nd edition), 170 p.

Lakoff, G., and Johnson, M., 1999, Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought: Basic Books, New York, 624 p.

Lee, N.Y.L., and Johnson-Laird, P.N., 2005, Synthetic reasoning and reverse engineering of Boolean Circuits, *in*, Mahwah, N.J., ed., Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society: Stresa, Italy, p. 1260-1265.

Leeds-Hurwitz, W. 1993, Semiotics and Communication: Signs, Codes, Cultures: Lawrence Erlbaum Associates, 222 p.

Luger, G.F., 2002, Artificial Intelligence: Structures and Strategies for Complex Problem Solving (4th ed.): Addison-Wesley, 856 p.

Miall, A.D., 2004, Empiricism and model building in stratigraphy: The historical roots of present-day practices: Stratigraphy, v.1, no.1, p. 3-25.

Miall, A.D., and Miall, C.E., 2004, Empiricism and model-building in stratigraphy: Around the hermeneutic circle in the pursuit of stratigraphic correlation: Stratigraphy, v. 1, no. 1, p. 27-46.

Nikolopoulos, C., 1997, Expert Systems - Introduction to First and Second Generation and Hybrid Knowledge Based Systems: CRC Press, 331 p.

Ogden, C.K., and Richards, I.A., 1923, The Meaning of Meaning: A Study of Language Upon Thought and the Science of Symbolism: New York, Harcourt, Brace & Company, Inc., 544p.

Parcell, W.C., and Williams, M.K., 2005, Mixed-sediment deposition in a retro-arc foreland basin: Lower Ellis Group (M. Jurassic), Wyoming and Montana, U.S.A.: Sedimentary Geology, v.177, n. 3-4, p. 175-194.

Primack, J.R. and Abrams, N.E., 2006, The View from the Center of the Universe: Discovering Our Extraordinary Place in the Cosmos, Riverhead, 400 p.

Raab, T. and Frodeman, R., 2002, What is it like to be a geologist? A phenomenology of geology and its epistemological implications: Philosophy & Geography, v. 5, no. 1, 69-81.