Analyzing the Psychological and Social Contents of Evidence—Experimental Comparison between Guessing, Naturalistic Observation, and Systematic Analysis

Haas, Henriette S ; Pisarzewska-Fuerst, Maja ; Tönz, Patrick ; Gubser-Ernst, Jutta

Abstract: To improve inferences about psychological and social evidence contained in pictures and texts, a five-step algorithm—Systematic Analysis (SA)—was devised. It combines basic principles of interpretation in forensic science, providing a comprehensive record of signs of evidence. Criminal justice professionals evaluated the usefulness of SA. Effects of applying SA were tested experimentally with 41 subjects, compared to 39 subjects observing naturally (naturalistic observation) and 47 subjects guessing intuitively intuitive guessing group. After being trained in SA, prosecutors and police detectives (N = 217) attributed it a good usefulness for criminal investigation. Subjects (graduate students) using SA found significantly more details about four test cases than those observing naturally (Cohen's d = 0.58). Subjects who learned SA well abducted significantly better hypotheses than those who observed naturally or who guessed intuitively. Internal validity of SA was a = 0.74. Applying SA improved observation significantly and reduced confirmation bias.

DOI: https://doi.org/10.1111/1556-4029.12703

Other titles: Systematisches Beobachten als Methode zur Analyse von psycho-sozialen Indizien in der Kriminalistik

Originally published at:
Haas, Henriette S; Pisarzewska-Fuerst, Maja; Tönz, Patrick; Gubser-Ernst, Jutta (2015). Analyzing the Psychological and Social Contents of Evidence—Experimental Comparison between Guessing, Naturalistic Observation, and Systematic Analysis. Journal of Forensic Sciences, 60(3):659-668.
DOI: https://doi.org/10.1111/1556-4029.12703

JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, 2015
doi: 10.1111/1556-4029.12703
Available online at: onlinelibrary.wiley.com

PAPER

PSYCHIATRY & BEHAVIORAL SCIENCES

Henriette S. Haas,[1] Ph.D.; Maja Pisarzewska Fuerst,[1] M.Sc.; Patrick Tönz,[1] M.Sc.; and Jutta Gubser-Ernst,[2] B.Sc.

Analyzing the Psychological and Social Contents of Evidence—Experimental Comparison between Guessing, Naturalistic Observation, and Systematic Analysis*

ABSTRACT: To improve inferences about psychological and social evidence contained in pictures and texts, a five-step algorithm—Systematic Analysis (SA)—was devised. It combines basic principles of interpretation in forensic science, providing a comprehensive record of signs of evidence. Criminal justice professionals evaluated the usefulness of SA. Effects of applying SA were tested experimentally with 41 subjects, compared to 39 subjects observing naturally (naturalistic observation) and 47 subjects guessing intuitively intuitive guessing group. After being trained in SA, prosecutors and police detectives (N = 217) attributed it a good usefulness for criminal investigation. Subjects (graduate students) using SA found significantly more details about four test cases than those observing naturally (*Cohen's d* = 0.58). Subjects who learned SA well abducted significantly better hypotheses than those who observed naturally or who guessed intuitively. Internal validity of SA was α = 0.74. Applying SA improved observation significantly and reduced confirmation bias.

KEYWORDS: forensic science, systematic analysis, naturalistic observation, intuitive guessing, abduction of hypotheses, psychosocial evidence, confirmation bias, experiment

In most criminal cases, social and psychological facts constitute an important part of the evidence: Written documents (letters or drawings), recorded statements, and crime scene photographs offer clues about the culprits' and witnesses' minds and personalities, namely memories, motives, plans, cognitions, and emotions as well as roles, relationships, proceedings, and capacities. Because only the mind's output can be observed, we must analyze (properly recorded) behavior and language and draw inferences based upon them.

The observation and interpretation of psychological and social aspects of evidence is an amorphous problem: It has not (yet) been clearly defined where to begin, how to proceed, and when the task is accomplished. In the absence of an established procedure to analyze the psychosocial contents of text and pictures, materials are often appraised by normal reading and then naively compared to the rest of the evidence. Countless studies in forensics (e.g., 1,2,3) have shown that under these circumstances, the interpretation of the evidence is prone to be biased. "Confirmation bias is a proclivity to search for or interpret additional information to confirm beliefs and to steer clear of information that may disagree with those prior beliefs [. . .]." (1).

An analytic approach based on generally accepted scientific principles should help to avoid considering only salient aspects of the observandum and falling prey to observer biases. Then again, some researchers in the decision-making field (4) advocate intuition over analytical thinking. They believe that accounting for too much information at a time can cause a cognitive overload and thus be counterproductive to good abduction. Behind this controversy is the information-processing predicament (2): "[. . .] the human brain has two distinct decision-making systems. The first is more analytic, rational, controllable and objective, whereas the second is more experiential and relatively independent of language, but is much faster."

Epistemological Principles of Systematic Analysis

To assist criminal investigators in unresolved cases, a method to improve the interpretation of social and psychological aspects revealed by the evidence was devised (5). Five rules, each one an undisputed principle in the scientific community, were organized into the algorithm of systematic analysis (SA) so as to observe questioned objects, pictures, and text from a social sciences point of view.

In everyday forensic practice, SA can be applied to anonymous letters and threats in order to establish an offender profile (6–8). Other fields of application are the interpretation of suicide notes, diaries, and last wills, of 911 calls, as well as of websites (9), drawings (10), plans, and photographs made by suspects or

[1]University of Zürich, Zürich ZH 8050, Switzerland.
[2]University of Basel, 4003 Basel, Switzerland.

*Supported by a research grant by the Foundation "Suzanne and Hans Biäsch for the promotion of applied psychology" (Stiftung Suzanne und Hans Biäsch zur Förderung der angewandten Psychologie, c/o Herr Prof. Dr. Ch. Steinebach, ZHAW Zürcher Hochschule für Angewandte Wissen-schaften, Departement Angewandte Psychologie, Minervastr. 30, Postfach, CH-8032 Zürich.

Received 6 Aug. 2013; and in revised form 26 Jan. 2014; accepted 20 April 2014.

witnesses, and crime scene photographs. Furthermore, analyzing recordings of interviews can help the interrogator to design more specific questions for follow-up interviews (7). We also recommend the method to describe crucial pieces of evidence about mens rea, such as letters, threats, or drawings in a murder charge (10). A context-blind expert can analyze these separately, to avoid a confirmation bias or the impression of a confirmation bias in the interpretation of the elements of mens rea.

First Principle: Using Schemata to Improve Perception

Theories of perception (11,12) assume that perception is a cognitive activity of integrating and comparing stimuli with mental schemata about the world. They reflect a structural understanding of the domain, thus reducing distraction by surface details and providing substantial information about the object in question. As a first rule, appropriate schemata of the observandum should be searched to insure that no important piece of information remain unacknowledged. To detect subtle differences and similarities, some schemata must be put side by side with the object of observation; for example, for the analysis of anonymous letters, the address on the envelope should be put physically next to the one in the phone book. The same is necessary for the comparison of addresses found on letters presumed to stem from the same author. When analyzing interview protocols, the questions form the schemata for all of the interviewee's subsequent answers. Provided that they have never been used in a preceding question before, all newly introduced words can be attributed unequivocally to the interviewee's mind (10).

Some schemata consist of ideal or statistical norms, of standards and plans; others consist of typologies of deviance (e.g., empirical studies about extortion letters). The comparison with a schema can do more than attract attention to details otherwise neglected; it may lead to new, unknown, or atypical aspects of the observandum.

The relevance of schemata to improve problem-solving skills was confirmed by several empirical studies (13–15).

Summing up Rule I: Compare the observandum with appropriate schemata, standards, or norms.

Second Principle: Recording Signs

The next step is to observe and describe all stimuli offered by the material. The smallest unit of perception is a *sign*, an entity with two inseparable faces, its form, and its contents. Peirce (16) defined three types of signs: indexes, icons, and symbols. An *index* is what we normally understand by sign of evidence, "it refers to the object that it denotes by virtue of being affected by that object". An *icon* (i.e., normally a picture) is a sign imitating the object which it stands for. A *symbol* is a sign referring to the object that it denotes by virtue of convention, tradition, or law. Observing the formal aspects of a text involves a description of its grammatical characteristics, of its choice of words such as professional languages, regional dialects, and finally elements of style (e.g., denials, use of passive voice). This part of the analysis can be enhanced using linguistic tools such as the LIWC software (17) and comparing the results with those from psycholinguistic studies (e.g., 18).

Empirical studies about the perception of form versus contents can be found in the fields of visual perception (19) and linguistics. Studying the recognition of formal syntactical structures, experiments showed that adults dismiss formal anomalies in texts more readily than contents anomalies (20).

Summing up Rule II: Describe the form of a sign not only its presumed meanings.

The nature of the sign has another procedural impact on scientific observation. A sign is to be understood as a coded signal within a specific social, cultural, and individual context (21). Consequently, signs should be interpreted within their context, which leads to the third rule.

Third Principle: Comprehensiveness and Structure

Unfortunately, we often do not know the entire context of a given case. The set of circumstances must also be reconstructed. Even if almost all cases involving living beings present complex open systems, we should strive for comprehensiveness. So reference classes should be chosen on the basis of all relevant knowledge available prior to the explanandum (22). To get as comprehensive record of all relevant signs as is possible, observers must apply the previously found schemata. They determine the underlying *structures* of the objects of the inquiry (e.g., photographs, drawings, letters, statements, web pages).

The third principle has been tested experimentally in the field of decision-making (23): When people organize their behavior into subgoals by segmenting their tasks according to a hierarchical structure, they consistently learn and observe better than when they do not. Observing each component of a structure is quite the same as segmenting behavior into subgoals.

When it comes to applying the rules II and III, the distinction between form and contents is relative within several layers of classification. For the purpose of analyzing a letter, we can list a relative relationship between form and contents in five different structures: (1) materials (e.g., DNA, ink . . .), (2) graphics of the document, (3) aspects of language, (4) elements of the letter type, and (5) actors and themes in the text (Table 1). Each of these structures is more formal than its neighbor on the right and more content oriented than its neighbor on the left. Observing all those structures of a letter (and not only structure 1) provides insight into an author's personality, his motives, skills, roles, and social environment. Observing actors includes identifying all grammatical actors (also pronouns and hidden actors in passive voice verbs or in nouns such as "a phone call"), but also modal verbs (want, must, may, need, can) and other latent pillars of meaning. These point toward Freudian slips hidden in a text.

For the analysis of pictures, all icons need to be translated into text to improve observation in criminal investigation (24).

Summing up Rule III: Observe each component of the object's structure as determined by the models.

Fourth Principle: Checking for Incoherence

The inventory of signs resulting from applying rules I to III might not be coherent. Anomalies, contradictions, and inconsistencies may appear. According to Aristotle's *principle of non-contradiction*, any valid hypothesis concerning the evidence must resolve such inconsistencies (25).

Empirical studies show that human propensity to overlook incoherence in pictures and in text is considerable. In an experiment on visual perception (19), it was found that people are quite insensitive to inconsistencies of illumination direction, even when they are specifically told to look out for them. Another study addressed the detection of incoherence in text (26): "the ability to detect inconsistencies or contradictions is an important aspect of comprehension. Indeed, a failure to do so may lead to unwarranted acceptance of a set of erroneous arguments and assumptions." It was

TABLE 1—*Checklist for Systematic Analysis (2014 version): follow the rules no matter whether you expect a promising result or not.*

Rule I. Find Schemata and Models for the Object of Observation	
Written Documents and Texts	Drawings and Photographs
Addresses on Internet and phone book Business or private letter (if necessary use culturally determined models) For other documents, use official standards Recorded interviews: text of questions figures as model for subsequent answers Schemata of deviance: empirical studies about extortion, threat, and suicide letters, about 911 calls	Drawings: made by children, artists, cavemen, scientific drawings, and drawing instructions Photographs: Several photographs of a case serve as models for each other Photographs: compare the conformity of human or animal anatomy and actions to "actions" visible in the photograph Schemata of deviance: drawings by mentally ill, drugged, or intoxicated persons, photographs and drawings by sex offenders
Rule II. Observe Formal Aspects of the Signs	
Written Documents and Texts	Drawings and Photographs
Documents: observe graphics, that is, format, layout, margins, edges, fonts/handwriting, representation of numbers, horizontal lines versus nonaligned text, use of icons, stains Text: observe language Punctuation, spelling Sentence structure (subordinate clauses, questions, denials, passive voice, etc.) Unusual words Style: local dialects, professional language, slang, malingering a foreign language speaker	Drawings: Two-dimensional versus three-dimensional (perspective), b&w versus color, secure/insecure hand, text, and figures in a drawing Drawings: observe proportions Photographs: light and reflections, dynamic versus static, undisturbed versus disturbed, alive versus dead Time if relevant: sequence, time of day, or season Space if relevant: relative position, entry versus exit, geography, cardinal directions
Rule III. Dissect the Object into Its Structural Components According to the Relevant Schemata	
Written Documents and Texts	Drawings and Photographs
Letters: observe the elements, that is, return address, inside address, city and date, salutation (according to purpose and culture), main body (according to purpose), closing, signature Texts: List all actors (humans, pronouns, institutions, animals), also hidden subjects (lack of commitment from speaker). Anonymous writers are often on this list (not only as "I" or "we"). List relevant themes, list terms used for weapons, threats, money, vehicles, drugs. Observe modal verbs (who may "want", who must "have to do things"?). List actions, cognitions, and emotions	Describe every component of the picture Specify human actors, objects, and animals Specify foreground versus background Describe potential human actions Describe potential natural influences Describe potential technological influences
Rule IV. Note Inconsistencies, Anomalies, and Contradictions	
Written Documents and Texts	Drawings and Photographs
Between one sentence and another Between the grammatical construction on the one hand and the meaning of words or sentences on the other hand Evolution: Observe how the same actor, the same object, or the same theme changes names along the text (in list of rule III) Check for contradictions between text and pictures	Between one picture and the other Within the components of a picture Within the drawing style, photographic conditions Contradicting human anatomy or human actions with the "actions" in the picture
Rule V. List What is Missing or Superfluous	
Missing or superfluous formal elements or signs see list for rule II Missing or superfluous components or aspects see list for rule III	

VI. Make a hypothesis check with a three-column table (for each relevant hypothesis)

Signs that are contradicting H0	*Signs that are inconclusive with respect to H0*	*Signs that are compatible with H0*

found that subjects who were instructed about the presence of contradictions doubled the number of identifications of them, but they also doubled the number of false positives. Testing children's failure to notice inconsistencies in text (27) researchers found that detection failures were related to poor recall of incoherence, caused by the inability to mentally represent inconsistencies. These results lead us back to the necessity of supporting memory by consulting appropriate schemata.

Summing up Rule IV: Detect inconsistencies, anomalies, and contradictions.

Fifth Principle: Recording Missing Elements

The unconscious assumption that *what you see is all there is* constitutes a main source for errors of judgment (28). Scientists should make missing signs explicit in all inquiries. The comparison of the evidence with the structures of the relevant schemata facilitates the awareness of what should or could be there, yet is not. The importance of *missing data* is basic to probabilistic reasoning (29,30), which might be applied to the material in a later step to resolve questions of proof. Then again, we want to notice apparently superfluous details (such as rejected valuables in a burglary or such as elements of an extortion letter that are not necessary to accomplish the crime), because they can also reveal important information.

Two studies (31,32) tested undergraduates' ability to identify sufficient, missing, and irrelevant information in math problems. Identifying those problems with sufficient information was the task best solved with a standardized score of $M = 0.84$ (SD = 0.22), while the discovery of missing information and that of irrelevant information received much lower scores with $M = 0.35$ (SD = 0.25) and $M = 0.24$ (SD = 0.23), respectively (31). Then, the perception of both contradictions and missing information within math problems and brainteasers was also examined (33): Undergraduates received a booklet containing tasks to solve. Some of those were in the original version, whereas others had been edited to contain contradictions or had some information deleted. The doctored problems were technically impossible to solve except by generating inferences. Subjects (not knowing this) were encouraged to formulate questions after having read a task. The results showed that the chances that the subjects did indeed ask questions (because they had spontaneously detected the insolvability of a task) were low, namely LR = 0.5 (SD = 0.35) for the missing information and LR = 0.31 (SD = 0.32) for the contradictions.

Summing up Rule V: Take note of all missing and superfluous elements or signs.

Applying the Algorithm of the Five Rules in Practice

In many criminal investigations, the first hours and days count most (34). Under the pressure to solve a major case, it may be difficult to figure out all potential schemata with their different structures. Therefore, we created a one-page checklist how to apply the method on pictures and on text (Table 1).

One of the test cases in the experiment shall illustrate the procedure. As part of a scientific study (35) in 1959, an adult male in full physical and mental health drew the sketch of an elephant and himself next to it (Fig. 1). The riddle to solve was: "Why would somebody draw like this?"

FIG. 1—*Drawing of an elephant (35).*

Rule I: We can use a scientific picture of an elephant as normative schema, but we will also consider other drawings such as children's or cave men's. In the drawing, we should compare the human and the animal figures as models to each other. Rule II, describing formal aspects of the observandum: The questioned drawing is a black and white pencil sketch, done with an insecure shaky hand. It lacks any notion of depth or perspective: There is neither foreground nor background. Nonetheless, proportions between man and elephant and between the animal's torso and its legs are correct. Rule III: All components of the icons need to be described according to the logic of the drawing (here: anatomy). If this is impossible, each sector of the paper can be numbered and described. The questioned elephant has four legs drawn like sticks, a large torso, a tail, and a trunk, but no head and no feet. The man has a head and a torso, whereas his limbs are too short, almost missing. His face is empty. Furthermore, the limbs of the elephant connect to its body and the head of the man connects to his body. Rule IV: The drawing contradicts anatomy in a grotesque way. Comparing it to cave men's drawings shows that they knew the anatomy of animals quite well and they also drew with a secure hand despite their primitive utensils. Contradicting the drawing's general awkwardness, the proportions of the figures seem just about right. Rule V: The most bizarre characteristic is that the elephant has no head at all (thus no ears, tusks, mouth, eyes) and that the man's face lacks eyes, nose, and mouth. A comparison of these features with children's drawings, where the head is the most prominent characteristic of all, reveals an unusual situation.

Abducting Hypotheses

Based on the inventory of signs, some hypotheses spring to the observers' minds. *Abduction*, according to Peirce (36), is the creative process of adopting an explanatory hypothesis that fits the data in retrospect. It is the third mode of logical inference besides deduction and induction. Every criminalist knows that diligent observation does not automatically lead to the abduction of the correct hypothesis. Most cases being open systems, there can be several plausible explanations fitting the available facts: "The identification of pattern in crime investigation may perhaps be defined simply as the identification of a deterministic sequence in a series of apparently chance events." (37).

At this stage of the inquiry, one may need to review the evidence in a re-iterative process, when the obtained results yield new insights and offer new ideas for consulting yet another model.

In the case of the elephant drawing, the hypothesis of drug intoxication might have sprung to our mind, so we might want to search for drawings made under the influence of drugs. Another explanation could be that the author has never seen an elephant in his life. Here, we could consider a sculpture made by a medieval artist who has obviously heard about elephants but never seen them (38). This sculpture is quite the opposite of the questioned drawing because it contains all the anatomical parts of an elephant but in distorted proportions. Other hypotheses are as follows: Was the man blindfolded? Was he drawing with his subdominant hand? Was he an artist? Was he raised in seclusion or in the wilderness? Was he under stress? Did he sabotage research? and Was he malingering?

Comparing Different Hypotheses, Checking for Validity

Finally, the most likely among different explanations must be chosen. For this purpose, all signs of evidence can be recorded within a three-column table (Table 1): those contradicting a specific hypothesis, those being inconclusive, and those that fit. The collection of evidence must now be criticized from all angles. To avoid a fragmentation of the evidence and the resulting fruitless discussions or erroneous decisions, we must be transparent about any decisions to include or exclude signs. Only a comprehensive table of all evidence can ensure falsifiability of the favored hypothesis (39). The table also permits to move signs from one column to another to follow up on the effect of temporary exclusion or inclusion of certain signs on the overall picture. Last but not least, a three-column validity check provides a record explaining a decision that seems to be the most reasonable in light of the available evidence—in case it must be defended against media or political pressure or in case some surprising developments in the investigation should later occur.

For the elephant drawing, the blindfold hypothesis must be rejected because it cannot explain why the legs are connected to the torso. The hypotheses of drawing with the wrong hand, an upbringing under seclusion, or in the wilderness cannot explain the missing head and face. The hypotheses of psychopathology (e.g., autism), neurological disorders (e.g., neglect, face-blindness), or severe visual impairment contradict the initial information, the fact that the man was physically and mentally sane at the time. Only the hypotheses of sabotage, artistry, malingering, and drug intoxication can be retained as plausible in light of the available evidence. The solution to the riddle is hard to find, because it is such an unusual scenario: The author of the drawing had been congenitally blind when a new type of surgery was performed on him, which gave him full vision. A team of cognitive psychologists had visited him in the hospital and had him draw things he had never seen, before showing them to him (35).

Experts' Opinion about the Algorithm

To determine the usefulness of the five rules for the criminal investigation, we analyzed anonymous evaluations of eleven 1-day courses in continuing education given to state prosecutors and police detectives ($N = 217$ participants since 2008). They graded the training in the five rules of analysis and hypotheses abduction on a scale from 1 (very poor) to 6 (very good).

The 43 prosecutors were 58% males and 42% females with an average age of $M = 38.5$ years (SD = 8.6) and a professional experience in criminal investigation of $M = 9.7$ years (SD = 8.3). They appraised the "usefulness" with a mean grade of $M = 5.6$ (SD = 0.6) and the "intelligibility" with a mean of $M = 5.7$ (SD = 0.5).

Chiefs of Police of eight local police corps had declared the participation of courses in SA mandatory for all their detectives. So this sample covers entire cohorts and includes the whole range of motivation, professional experience, and skill. The 174 detectives (91.3% males and 8.7% females) averaged $M = 41.8$ years of age (SD = 8.1) and had $M = 16.8$ years (SD = 8.9) professional experience in criminal investigation. They graded the usefulness of the method with a mean of $M = 4.7$ (SD = 0.7), and they valued it as a good critical thinking tool with a grade of $M = 5.0$ (SD = 0.6).

Experimental Procedure

Participants

Participants of the experiment were graduates regularly enrolled in classes of forensic psychology and criminology. Different semesters built three quasi-random samples. The "SA" group was instructed with the five rules and practiced them; it consisted of $n = 41$ participants (33 females, six males, 2 missing). They were told that they would learn criminal profiling during the experiment. Another $n = 39$ students (33 females, four males, 2 missing) were in the "naturalistic observation" group (NO). They were told that they would get the opportunity to "learn by doing casework" in an experiment (40). Those 80 students stayed for all three sessions, completing the whole test series and filling out the questionnaires, while another 15 students dropped out before completing the experiment. The intuitive guessing group (IG) consisted of 47 students (32 females, 14 males, 1 missing). They were instructed that the goal was to compare intuition to careful observation and description in casework. They completed the tests cases by pure guessing within one session. All 127 students were white Europeans and proficient in the language of instruction. The mean age of the SA group was $M = 25.9$ years (SD = 5.6), that of the NO was $M = 26.1$ years (SD = 5.4), and that of the IG was $M = 26.8$ years (SD = 5.6). Age differences were not significant. The participants' sex had no significant effect on the outcome variables.

Experimental Setting

The test cases (Table 2) were selected from several dozen. They had to have a proven ground truth and offer the opportunity to apply all five rules. We wanted to present a variety of themes and causes.

TABLE 2—*Test cases presented in the experiment.*

Test Case	Description
A. Broken window	Photographs of a broken window (star-shaped hole) next to the entrance of a suburban home. What happened? The ground truth: No human mind was involved. A hail storm had hit the house. Clues to detect: No object having caused the hole was visible. The broken window was barred and provided no opportunity to enter the house, while a plant on the windowsill was is shreds
B. Elephant drawing	The elephant drawing (35) was done as part of a scientific study. Why would a mentally and physically healthy man draw like this? The ground truth: A man recovering from congenital blindness (by surgery), who had never seen an elephant before, had drawn the picture. Clues to detect: His insecure drawing hand, the animal's head is missing, while the limbs are connected to the body and the figures are in adequate proportions
C. Murder scene	Photographs of a crime scene (44) and a description of the events leading up to the discovery of a dead female in her apartment were presented. Who was the main suspect? The ground truth: The murderer was the husband who had staged the crime scene to appear as a burglary. Clues to detect: The inconsistency of the manipulated crime scene combined with the strange behavior of the husband when the crime was discovered
D. Citizen's letter	An individual mentioned in a homeland security report on extremism, as being a leading negationist (Holocaust denier), wrote a letter of protest to the Prime Minister. Was he a negationist or not? The ground truth: The author was a notorious negationist convicted for organizing conferences and propagating racism. Clues to detect: Numerous and superfluous grammatical negations, a total absence of terms specifying historical facts (e.g., "genocide"), combined with a reversal of the meaning of the word "negationist"

Concerning the final hypothesis resulting from the observations, we chose two ways to request it on the answering sheets. For the two criminal cases, the *murder scene* and the *citizen's letter*, we presented a choice of several hypotheses. For the two riddles, the *broken window* and the *elephant drawing*, subjects depended entirely on their imagination as we just left open space for them to state their best hypothesis. Furthermore, we instructed them to define only one final hypothesis.

To avoid that the experimental effect be flawed by inevitable differences in the test cases, the participants of each group were randomly assigned to subgroups. Subgroup 1, representing half of the class, had to treat cases A and D during the first test, whereas subgroup 2 had to treat cases B and C at first. During the second test, the participants received the inverse pair of cases to test their performance.

During the tests, participants were working on their own in a quiet classroom, seated with an empty chair between each other. They were told not to exchange any information about the test cases during the whole experiment. This examination-like procedure was strictly enforced.

The SA and the NO group both started with the first 1-h test about two cases to establish the baseline of their capacity to observe and interpret. They would carefully read the cases, study the pictures, take notes about their observations, and finally, state their best hypothesis. The participants in the observation groups (SA and NO) received lined sheets of paper with enough space to describe all their observations. The experimental group in SA then received 5 h of training, spread over the same and the next afternoon. It included a 90-min presentation explaining the rules, illustrated by analyzing the criminal manifesto of the anthrax attack (5,6,41). The students also practiced the systematic application of the rules with four exercises (a photograph, an anonymous threat letter, a drawing, and a newspaper interview with a terrorist suspect). Results of those exercises were discussed in plenum. For the second test, the SA group received the checklist (i.e., a previous version). During the same time span, the NO control group received presentations about different topics of forensic psychology not related to any of the rules of SA. During the third afternoon, each group was asked to take the second 1-h test on the inverse pair of cases.

The IG received the instruction to write down the first hypothesis springing to their mind immediately upon reading a case. This was practiced a few times. To ensure that all subjects spend the same amount of time reading, each case was read aloud slowly while the students would also follow it on paper. Then, they were given 60 sec to write down their best hypothesis about the case. The answer sheets left no space to put down any observations. After the first test, they changed rooms and did the inverse pair of test cases.

Operationalization of the Rules

The subjects' use of rules in the SA and NO groups was measured using the categories specified in the checklist (counting the number of schemata, formal aspects, components, inconsistencies, and missings according to the old version of the checklist). The overall wealth of signs that participants had observed was measured by taking the sum of observations according to all rules.

To measure the predictive validity of the three different approaches, we took the quality of the favored hypothesis. All those hypotheses, which were either plausible or a hit of the ground truth were considered to be logically *valid*. Finding a valid hypothesis involves two steps: first, the abduction of good ideas, and second, critical thinking whether the favored hypothesis is consistent with all signs of evidence. Those two additional reasoning modes and the instruction of inserting all signs of evidence into a three-column table were just mentioned in class and not practiced with the SA group.

Rating the Participants' Answers

All authors participated in the coding of the participants' answers. To ensure double-blinding, the control variable (first vs. second test) was covered on the subjects' response sheets. Unfortunately, it was sometimes possible to spot analyses carried out by trained participants, because they explicitly referred to the rules.

To rate the application of each rule, we established lists for all answers falling under a given rule. The attribution to rules I and III was less ambiguous than the one to the rules II, IV, and V. Certain observations could be rated under either of two of the latter rules. We coded such answers under one rule only (always the same for all subjects).

The logical validity of the subjects' favored hypotheses about a case was determined by the first and second co-author. Beside hits of the ground truth, we coded those hypotheses as a *plausible* that explained all the presented evidence of the case and contradicted none of it. All hypotheses that explained the facts only partially (e.g., the man has never seen an elephant), which contradicted some of the evidence (e.g., the man was blind) or which were too general to be refutable (e.g., the man was bad at drawing), were rated as "not valid".

To check the inter-rater consensus, the first, third, and forth co-authors rated the proficiency of the experimental group's use of the rules. The correct use of the rules on the test cases was graded on a 6-point scale (from 1 = "no use" to 6 = "excellent use"). When the joint application of all rules is considered, the correlation between different judgments about the cases treated by the trained subjects was *Cronbach's* $\alpha = 0.75$. Differences before and after learning the method showed no consistent picture. Obviously, a good agreement can be achieved on the total absence of guidance by a rule before it had been taught. Inter-rater disagreements pointed out some weaknesses of the checklist. Some subjects had found good proceedings not mentioned in the checklist.

Statistical Analyses

Statistical analyses were performed using SAS® 9.3 (PROC MEANS, PROC GLM, PROC ANOVA, PROC CORR, PROC FREQ, PROC UNIVARIATE SAS Institute Inc., Cary, NC, USA).

Results of the Experimental Testing

Effects of Learning the Rules

Not everybody who has been taught a method can integrate the new knowledge and apply it correctly, especially if there is no examination to enforce serious studying. Thus, we examined the difference between the NO and the SA groups, but we also considered the "good learners" among the latter ($n = 21$ or 51.2%). Those were defined (ex post) as the subjects who had—in both tasks of the second test—mentioned at least one schema. Between the "good learners" and the other 20 experimental subjects of the SA group, there were no significant differences during the first test: neither in the wealth of details observed, nor in the quality of their hypotheses.

Effect on the Wealth of Observations

We tested the external validity of the five rules by taking the mean intra-individual differences between the first and the second test of the SA and NO groups. Table 3 shows the improvements in the wealth of observations over two case analyses after the second test.

The NO group showed no significant progress, whereas the SA group increased the number of relevant observations during the second test significantly, effect size estimates being medium. Describing the structural components of an object (rule III) turned out to be the *default mode of naturalistic observation*, whereas the rest of the principles of SA must be taught and practiced. For the good learners, the effects of learning the five rules were very strong.

Predictive Validity in the Validity of Hypotheses

Does SA improve the quality of abducted hypotheses compared to NO and to IG? Being a creative task, the generation of well-fitting hypotheses cannot depend on exhaustive analysis alone. Creativity as an individual gift might explain an important part of the variance (42). Yet another (uncontrolled) influence on the quality of abduction may be general knowledge about the world and professional experience (4,11,28). In the experiment, such "general wisdom" concerned knowing how children or cavemen draw animals or knowing that many people tend to deceive others by omission rather than by overt lying. Contrary to analyses carried out in real life, the experimental subjects had to resort to their memories as models when it came to find schemata for drawings.

Finding out the ground truth was more difficult for the riddles, where it depended entirely on the subjects' imagination, than for the criminal cases, where a choice of multiple options was offered (Table 4). The SA group improved their hypotheses somewhat in the second test contrary to the NO group, but the effect size estimates were small. However, applying the five rules proficiently, just like the "good learners" had done, leads indeed to a significant increase of valid hypotheses. (Not shown in Table 4: For the IG group, we did the same comparison and found no significant improvement between the first and second test.)

Looking at absolute values in the quality of the abducted hypotheses (Table 5) between the three groups, we were quite surprised that IG on the spot worked just as well as (untrained) careful reading and describing the evidence (NO) for an hour per two cases. Thus, NO (following no specific rules) provided no added value over intuition after reading the case for the first time. We interpret this result as an expression of the confirmation bias. Maybe participants had formed an instant opinion about a case and had stuck with it throughout their observations. In the absence of an objective frame of reference, people tend to take their own first opinion as the reference and then reason backwards observing only those details confirming their initial hunch.

Applying the algorithm of the five rules did increase the quality of abducted hypotheses; especially the good learners scored more than 70% valid hypotheses compared to only about 40% for IG and NO.

Training in SA increased the odds to hit the ground truth by a LR = 1.55 compared to the results of untrained observers (with a *95% CI* of [0.93–2.60] and $p < 0.09$). Good learners' results showed an improvement by a LR = 2.24 to find out the truth of the case (with a *95% CI* of [1.16–4.35] and $p < 0.02$) in comparison with the results of untrained observers. The training also improved the abduction of a hypothesis that was at least plausible (instead of being false, illogical, incomplete, or too vague), even if it did not correspond to the ground truth. The odds to find a plausible (yet untrue) hypothesis by applying SA were increased by a LR = 2.25 (*95% CI of* [1.02–4.94], $p < 0.04$). For the good learners, it was a LR = 3.05 (*95% CI* of [1.22–7.60], $p < 0.01$).

Internal Validity of the Five Rules

What about the internal validity of the fives rules, do they form an algorithm, or are they just a loose bundle of recipes?

TABLE 3—*Intergroup comparison of the individual improvement in the wealth of observations during the second test.*

	Naturalistic Observation n = 39			Systematic Analysis n = 41 (Good Learners n = 21)			GLM (Unbalanced Data)			Effect Size Measures		
Operationalization	M	SD	95% CI	M	SD	95% CI	SS	$F_{1,76}$ ($F_{1,58}$)	$p<$	Partial η^2	Cohen's d	Effect size r
Rule I	−0.03	1.13	−0.39, 0.34	0.88	1.29	0.47, 1.28	16.3	11.0	0.001	0.124	0.744	0.349
Schemata used				*(1.52)	*(1.36)	*(0.90, 2.14)	–*	–*	–*	–*	–*	–*
Rule II	−0.33	2.73	−1.22, 0.55	1.54	2.43	0.77, 2.30	69.9	10.5	0.002	0.119	0.724	0.340
Formal aspects observed				(2.19)	(2.69)	(0.96, 3.41)	(86.9)	(11.8)	(0.001)	(0.170)	(0.926)	(0.420)
Rule III	0.21	5.41	−1.55, 1.96	0.61	5.94	−1.27, 2.48	3.3	0.1	0.751	0.001	0.071	0.036
Components observed				(2.38)	(6.07)	(−0.38, 5.14)	(64.6)	(2.0)	(0.160)	(0.034)	(0.382)	(0.188)
Rule IV	0.10	2.77	−0.80, 1.00	1.27	2.09	0.61, 1.93	27.2	4.6	0.036	0.036	0.475	0.231
Incoherences observed				(1.29)	(2.03)	(0.36, 2.21)	(19.1)	(3.0)	(0.091)	(0.049)	(0.505)	(0.245)
Rule V	−0.10	1.70	−0.65, 0.45	1.24	2.75	0.37, 2.11	36.2	6.8	0.011	0.011	0.588	0.282
Missings observed				(1.90)	(2.91)	(0.58, 3.23)	(55.0)	(11.4)	(0.001)	(0.165)	(0.841)	(0.388)
Total wealth of observation (all five rules combined)	−0.15	7.02	−2.43, 2.12	5.54	11.97	1.76, 9.32	647.2	6.6	0.012	0.078	0.580	0.278
				(9.29)	(12.81)	(3.46, 15.12)	(1216.3)	(13.7)	(0.001)	(0.191)	(0.921)	(0.418)

N = 80 subjects (observations over two analyzed cases per test).

*Tautological as "good learners" are defined by having used at least one schema for analyzing each of their cases in the second test.

TABLE 4—*Intergroup comparison of the individual improvement in the quality of the abducted hypothesis after the second test.*

	Naturalistic Observation $n = 39$			Systematic Analysis $n = 41$ (Good Learners $n = 21$)			GLM (Unbalanced Data)			Effect Size Measures		
	M	SD	95% CI	*M*	SD	95% CI	SS	$F_{1,78}$ ($F_{1,58}$)	*p*<	Partial η^2	Cohen's *d*	Effect Size *r*
Hits of the ground truth only	0.15	0.59	−0.04, 0.34	0.17 (0.38)	0.70 (0.50)	−0.05, 0.39 (0.15, 0.61)	0.01 (0.70)	0.01 (2.26)	0.908 (0.138)	0.000 (0.038)	0.026 (0.418)	0.013 (0.204)
Plausible hypotheses only	−0.06	0.54	−0.24, 0.13	0.20 (0.33)	0.58 (0.69)	0.00, 0.40 (−0.01, 0.67)	1.16 (1.81)	3.66 (5.71)	0.060 (0.027)	0.051 (0.092)	0.457 (0.632)	0.223 (0.301)
Valid hypotheses (plausible or hits)	0.08	0.70	−0.15, 0.30	0.37 (0.67)	0.83 (0.73)	0.10, 0.63 (0.33, 1.00)	1.69 (4.75)	2.81 (9.35)	0.098 (0.003)	0.035 (0.139)	0.376 (0.822)	0.185 (0.380)

$N = 80$ subjects (means over two cases per test, only one hypothesis was allowed par case), hypotheses of the two subjects who had known the elephant case were counted as not valid.

TABLE 5—*Quality and specificity of the abducted hypotheses according to the procedure (Systematic analysis [SA] and intuitive guessing group [IG] compared to naturalistic observation [NO]).*

	1 Min Per Case			30 Min Per Case					
	IG			NO			SA		
Test Case	Hits	Plausible Hypothesis	Valid Hypothesis (Plausible or Hit)	Hits	Plausible Hypothesis	Valid Hypothesis (Plausible or Hit)	Hits	Plausible Hypothesis	Valid Hypothesis (Plausible or Hit)
Elephant drawing (open question)		*n = 46 cases*			*n = 57 cases*			n = 21 cases (all SA) *n = 13 cases (good learners)*	
	0.0^{ns}	13.0^{ns}	13.0^{ns}	0.0	14.0	14.0	0.0^{ns} 0.0^{ns}	$38.1^{\dagger}$ $53.9^{\ddagger}$	$38.1^{\dagger}$ $53.9^{\ddagger}$
Broken window (open question)		*n = 47 cases*			*n = 61 cases*			n = 19 cases (all SA) *n = 8 cases (good learners)*	
	4.3^{ns}	0.0^{ns}	4.3^{ns}	1.6	4.9	6.6	10.5* $25.0^{\ddagger}$	10.5^{ns} 12.5^{ns}	21.0* $37.5^{\ddagger}$
Murder scene (multiple choice)		*n = 47 cases*			*n = 58 cases*			n = 22 cases (all SA) *n = 13 cases (good learners)*	
	72.3^{ns}	2.1^{ns}	74.5^{ns}	58.6	3.5	62.1	72.7^{ns} $92.3^{\dagger}$	4.6^{ns} 0.0^{ns}	77.3^{ns} $92.3^{\dagger}$
Citizen's letter (multiple choice)		*n = 47 cases*			*n = 61 cases*			n = 19 cases (all SA) *n = 8 cases (good learners)*	
	76.6^{ns}	2.1^{ns}	78.7^{ns}	70.5	6.6	77.1	89.5* 100.0*	5.3^{ns} 0.0^{ns}	94.7* 100.0^{ns}
All four test cases together		*n = 187 cases*			*n = 237 cases*			n = 81 cases (all SA) *n = 42 cases (good learners)*	
	38.5^{ns}	4.3^{ns}	42.8^{ns}	32.9	7.2	40.1	43.2* $52.4^{\dagger}$	$14.8^{\dagger}$ $19.0^{\dagger}$	$58.0^{\ddagger}$ $71.4^{§}$

$N = 127$ subjects over 508 test cases.
Significance of the chi-square test compared to naturalistic observation: $*p < 0.1$, $^{\dagger}p < 0.05$, $^{\ddagger}p < 0.01$, $^{§}p < 0.001$.

The mechanics of the rules (Fig. 2) show that they are correlated in a meaningful way. Finding schemata (rule I) associates only with the observation of the underlying formal and contents structures of the observandum, that is, rules II and III combined. The latter rules do not depend on each other, nor do rules IV and V. The role of rule III as the default mode of NO is also reproduced within the mechanics of the algorithm: Contrary to rule II, there is no significant correlation to the use of schemata. Then again, a rich record of components according to rule III provides a better recognition of anomalies and inconsistencies (rule IV). Finally, there is a link between the observation of formal elements and missings. Internal validity between the first part of the analysis (rules I to III) and the second part tested with a Cronbach's $\alpha = 0.74$ for standardized variables (and $\alpha = 0.67$ for raw variables). Thus, the five rules are built upon each other and not redundant.

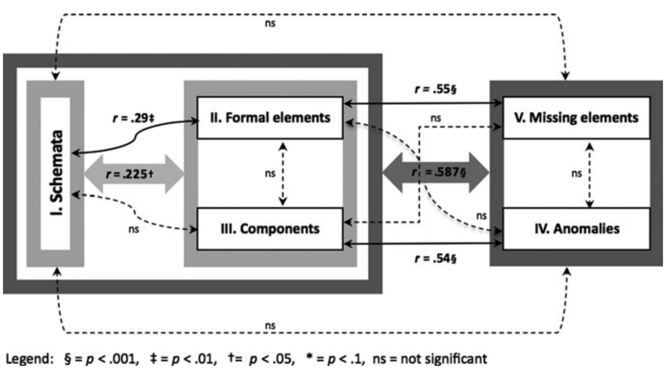

FIG. 2—*Internal validity of the five rules of systematic analysis,* N = *82 case analyses carried out by the 41 trained experimental subjects.*

Interpretation of the Experimental Results

The answering sheets of the participants observing naturally (NO) made it obvious that looking for psychosocial facts in the

TABLE 6—*Synopsis of advantages and disadvantages of the three procedures (IG, NO, SA).*

	Procedure		
	IG	NO	SA
Procedure is quick	Yes	No	No
Procedure is low cost	Yes	No	No
No need for particular intellectual effort or increased concentration	Yes	Yes	No
No need for special instruction	Yes	Yes	No
Procedure favors problem-solving and critical thinking skills in general	No	No	Yes
Procedure provides a neutral reference frame for the object of observation (unrelated to the case's context)	No	No	Yes
Procedure leads to a greater wealth of observed details	No	No	Yes
Procedure leads to the abduction of better hypotheses	No	No	Yes
Procedure reduces confirmation bias	No	No	Yes
Procedure provides an overview of all signs of evidence and their interpretation so that results or decisions can be criticized from all angles (refutability)	No	No	Yes
Procedure can increase the chance to abduct plausible yet mistaken hypotheses	No	No	Yes

IG, intuitive guessing; NO, naturalistic observation; SA, Systematic analysis.

evidence is an ill-defined task. Already during the experiment, we could directly observe that most untrained subjects were lost; they did not know where to begin and many had finished the tasks before the end of the test period (but they were not allowed to leave or to talk). The trained participants on the other hand were more focused and many would have preferred more time to solve the tasks. Then again, the procedure demands a certain discipline, and so the benefit of applying the five rules on the quality of the hypotheses manifested itself more clearly in the "good learners".

While the SA training raised the likelihood to find the ground truth, it also raised the likelihood to abduct a plausible yet untrue hypothesis (instead of hypotheses clearly contradicting the evidence which were generated by subjects without training). Yet abductions of plausible but untrue hypotheses constitute somewhat of a quandary. While it is desirable to open up one's mind to as many plausible alternatives when treating unresolved cases, such hypotheses compete with the unknown ground truth of the case. The example of the elephant drawing illustrates the danger that professionals stop looking for alternatives after finding a first plausible explanation (e.g., drug intoxication). Some plausible but mistaken hypotheses can be quite convincing and constitute pitfalls for criminal investigation.

During the rating process, we realized that improvements must be made to the checklist, especially for the analysis of pictures. According to our results, all five rules build upon each other and are necessary steps in the process of analysis; none of them can be dismissed.

Applying this method yielded a more comprehensive picture of the evidence, which in turn helped to find better explanations for the facts. But given the fact that this experiment is only the first of a series to come and was performed by those involved in creating the method, its results need to be appraised with some caution. As the first author who conceived the five-step algorithm is also the one who taught it, and who taught "learning by doing" and "guessing by intuition," the results of the experiment might be influenced by an interpersonal expectancy effect (43). Another concern might be the fact that a perfect blind rating was not always possible (because subjects had revealed that they were working according to SA), while at the same time, the ratings of the subjects' questionnaires were not independent from the creation of the method. Furthermore, it could be criticized that the criterion for "good learners" had been stated ex post and not ex ante. Finally, the test cases had to be selected for solvability, thus providing the opportunity to obtain better results than what is to be expected otherwise.

All these issues of concern are weaknesses encountered typically when a method undergoes its first scientific testing. So if the present results raise hopes for applying systematically the scientific principles of forensic science to psychosocial contents of the evidence, they need to be reproduced by independent researchers before drawing definite conclusions. Furthermore, one could argue that this experiment does not demonstrate the effect of those five rules specifically. Maybe people pay more attention to detail when following any well-defined procedure rather than just observing naturally. Therefore, further experimenting is required to test effects of a placebo procedure and compare it with SA.

Nevertheless, the results of this first experiment fall in tune with all theories mentioned in the introduction and even resolve some controversy between them. Intuition—that is pure guessing (IG)—turned out to be better than its reputation (5) because NO was prone to be influenced by a confirmatory bias (1–3). Just like applying one or two of the rules showed improvements in problem solving (23,26,28,31,32), the SA algorithm as a whole did also increase the wealth of details observed and the validity of abductions made. Thus, its application decreased the risk to fall prey to biased perception, especially when the method was proficiently applied (2,23,28). Table 6 offers a synopsis over advantages and disadvantages of each one of the three approaches.

Finally, we must mention a general risk encountered in all methods used to abduct plausible hypotheses. Professionals without training in empirical research often attribute more predictive validity to the results of an analysis or an offender profile than they actually have (sometimes even mistaking them for facts). All reports based on applying methods to reveal psychosocial contents in criminal cases must therefore begin with some introductory remarks about the procedures' limitations, namely about the base rates of true hits and about the occurrence of plausible yet mistaken hypotheses.

Conclusion

According to this first experiment, the five general scientific principles united under the algorithm of SA provide a broader mind-set of what can be perceived about social and psychological aspects contained in the evidence and of what must be considered before an observer can abduct far-reaching hypotheses. Some other features of the method should yet be examined: What are potential pitfalls and sources of errors related to the application of this method and what is the inter-rater reliability of the five rules over many different cases?

The following points could be interesting for future experiments: As a consequence of the large part of unexplained variance in the quality of the abducted hypotheses, individual intelligence differences namely in deductive reasoning and creativity could be examined and compared to the success of analyzing test cases with different methods.

In prosecutors' and police detectives' opinion, the five-step procedure provides a useful tool for criminal investigation, helping them to decide which tracks to pursue and how to bring together all available signs of evidence in a falsifiable way. Yet an approval cannot replace actual testing. A similar experiment with experienced professionals is essential.

Acknowledgments

We want to express our thanks to René Hirsig, professor of methodology, to Brigitte Boothe, professor of clinical psychology, and to Thomas Hansjakob, JD, chief prosecutor of St. Gall, for their generous advice and support and to Marianne Pithon for her close reading of the manuscript.

References

1. Budowle B, Bottrell MC, Bunch SG, Fram R, Harrison D, Meagher S, et al. A perspective on errors, bias, and interpretation in the forensic sciences and direction for continuing advancement. J Forensic Sci 2009;54 (4):798–809.
2. Dror IE. A novel approach to minimize error in the medical domain: cognitive neuroscientific insights into training. Med Teach 2011;33:34–8.
3. Dror IE, Charlton D, Peron A. Contextual information renders experts vulnerable to making erroneous identifications. Forensic Sci Int 2006;156 (1):74–8.
4. Klein G. The power of intuition. New York, NY: Currency Books, 2003;21,271,284.
5. Haas H. Kriminalistischer Erkenntnisgewinn durch systematisches Beobachten. Kriminalistik 2003;57:93–100.
6. Haas H. Systematic observation as a tool in combating terrorism. In: Pieth M, Thelesklaf D, Ivory R, editors. Combating the financing of terrorism. Basle Institute on Governance. New York, NY: Peter Lang, 2009;59–93.
7. Haas H. Verlaufsanalysen von häuslicher Gewalt. In: Bundesministeriums für Justiz, editor. Lebensform Familie—Realität & Rechtsordnung. Österreichische RichterInnenwoche Laa an der Thaya 26. bis 30. Mai 2008. Schriftenreihe des BJ Band 141, Wien, A: Neuer Wissenschaftlicher Verlag, 2009;121–43.
8. Haas H. Analyse anonymer Schreiben unter Berücksichtigung von Gender-Aspekten. In: Steins G, editor. Handbuch Psychologie und Geschlechterforschung. Wiesbaden, D: VS Verlag für Sozialwissenschaften, 2010;249–70.
9. Renfer M, Haas H. Systematic analysis in counterterrorism: messages on an Islamist Internet-forum. Int J Intell CounterIntell 2008;21(2):314–36.
10. Haas H, Sutter L. Methoden der Feststellung von psychischen Tatsachen im Strafrecht. Aktuelle juristische Praxis (AJP)/Pratique Juridique Actuelle (PJA) 2013;11:1589–99.
11. Gregory RL. The Medawar lecture 2001. Knowledge for vision: vision for knowledge. Philos Trans R Soc Lond B Biol Sci 2005;360:1231–51.
12. Neisser U. Cognition and reality. San Francisco, CA: Freeman, 1976;7.
13. Gick ML, Holyoak KJ. Analogical problem solving. Cogn Psychol 1980;12:306–55.
14. Brown AL, Kane MJ. Preschool children can learn to transfer: learning to learn and learning from example. Cogn Psychol 1988;20:493–523.
15. Brown AL, Kane MJ, Long C. Analogical transfer in young children: analogies as tools for communication and exposition. Appl Cogn Psychol 1989;3:275–93.
16. Peirce CS. Elements of logic. In: Hartshorne C, Weiss P, editors. Collected papers of Charles Saunders Peirce, Vol. II. Cambridge, MA: Harvard Univ Press, 1978/1931;2.275.
17. Pennebaker JW, Booth RJ, Francis ME. Linguistic inquiry and word count: LIWC. Austin, TX: 2007; www.liwc.net (accessed December 30, 2013).
18. Argamon S, Koppel M, Pennebaker JW, Schler J. Automatically profiling the author of an anonymous text. Commun ACM 2009;52(2):119–23.
19. Ostrovsky Y, Cavanagh P, Sinha P. Perceiving illumination inconsistencies in scenes. Perception 2004;34:1301–14.
20. Braze D, Shankweiler DP, Ni W, Palumbo LC. Readers' eye movements distinguish anomalies of form and content. J Psycholinguist Res 2002;31 (1):25–44.
21. Eco U. Segno. Milano, I: Istituto Editoriale Internazionale, 1973;5.19.
22. Taroni F, Aitken C, Garbolino P, Biedermann A. Bayesian networks and probabilistic inference in forensic science. Chichester, West Sussex, U.K.: John Wiley & Sons, Ltd., 2006;25.
23. Hard BM, Lozano SC, Tversky B. Hierarchical encoding of behavior: translating perception into action. J Exp Psychol 2006;135 (4):588–608.
24. Oevermann U. Die Methode der Fallrekonstruktion in der Grundlagenforschung sowie der klinischen und pädagogischen Praxis. In: Kraimer K, editor. Die Fallrekonstruktion. Sinnverstehen in der sozialwissenschaftlichen Forschung. Frankfurt a. M., D: Suhrkamp, 2000;58–156.
25. Aristotle. Metaphysics. In: Barnes J, editor. The complete works of Aristotle: the revised Oxford translation, Vol. II. Princeton, NJ: Princeton University Press, n.d./1984;1552–728;1597.
26. Ellis HC, Ottaway SA, Varner LJ, Becker AS, Moore BA. Emotion, motivation, and text comprehension: the detection of contradictions in passages. J Exp Psychol 1997;126(2):131–46.
27. Vosniadou S, Pearson PD, Rogers T. What causes children's failures to detect inconsistencies in text? Representation versus comparison difficulties. J Educ Psychol 1988;80(1):27–39.
28. Kahneman D. Thinking, fast and slow. London, U.K.: Peguin Books, 2011;86.
29. Dodge Y. Analysis of experiments with missing data. Chichester/West Sussex, U.K.: John Wiley & Sons, 1985.
30. Lessler J, Kalsbeek W. Nonsampling error in surveys. New York, NY: John Wiley & Sons, 1992.
31. Birney DP, Fogarty GJ, Plank A. Assessing schematic knowledge of introductory probability theory. Instr Sci 2005;33(4):341–66.
32. Low R, Over R. Detection of missing and irrelevant information within algebraic story problems. Br J Educ Psychol 1989;59:296–305.
33. Graesser AC, McMahen CL. Anomalous information triggers questions when adults solve problems and comprehend stories. J Educ Psychol 1993;85:136–51.
34. Keppel R, Weis W. Time and distance as solvability factors in murder cases. J Forensic Sci 1994;39(2):386–400.
35. Gregory RL, Wallace JG. Recovery from early blindness—a case study. Experimental Psychology Society Monograph 2, 1963;31; www.richardgregory.org/papers/recovery_blind/contents.htm (accessed December 30, 2013).
36. Peirce CS. Pragmatism and pragmaticism. In: Hartshorne C, Weiss P, editors. Collected papers of Charles Saunders Peirce, Vol. V. Cambridge, MA: Harvard University Press, 1978/1931;5.171.
37. Kind S. The scientific investigation of crime. Manchester, U.K.: Forensic Science Services, 1987;43.
38. Wikimedia. Sculptures on the Basle Münster (CH). http://commons.wikimedia.org/wiki/File:Elefant_am_Basler_Muenster.jpg http://www.baslermuenster.ch/bauwerk/rundgang/aussenbau/chor (accessed December 27, 2013).
39. Russell B. The problems of philosophy. Oxford, U.K.: Oxford University Press, 1912;89.
40. Piserzewska M. Kriminalistisches denken [M.Sc. thesis]. Zürich: University of Zürich, 2012.
41. Federal Bureau of Investigation (FBI). Headline archives anthrax investigation–closing a chapter. Aug. 6, 2008; www.fbi.gov/page2/august08/amerithrax080608a.html (accessed December 30, 2013).
42. Carson SH, Peterson JB, Higgins DM. Decreased latent inhibition is associated with increased creative achievement in high-functioning individuals. J Pers Soc Psychol 2003;85(3):499–506.
43. Rosenthal R, Rubin DB. Interpersonal expectancy effects: the first 345 studies. Behav Brain Sci 1978;1(3):377–415.
44. Trapp C. Staging des Tatorts und andere Inszenierungen bei Tötungsdelikten. Zürich: Schulthess Verlag, 2012;13–25.

Additional information and reprint requests:
Henriette Haas, Ph.D.
Adjunct Professor of Forensic Psychology
Psychologisches Institut der Universität Zürich
Binzmühlestr. 14/1 Box 1
CH-8050 Zürich
Switzerland
E-mail: henriette.haas@psychologie.uzh.ch