Interpretation: Observer Effects

Introduction

The term observer effect generally describes circumstances in which the results of an observation are affected by the observer. In physics, the term has been used to describe circumstances in which the act of observation changes the phenomenon being observed, as where measurement of electrical current in a circuit changes the current flow. By contrast, in the social sciences and in forensic science, the term is used to describe circumstances in which the observer’s preconceptions or motives influence conclusions drawn from data [1]. The preconceptions and motives are thought to influence the perception and interpretation of the evidence rather than changing the evidence itself. Observer effects are sometimes also called examiner bias [2, 3] although it important to note that the “bias” entailed in the phenomenon may occur without the observer intending or even being aware of it [4, 5].

Observer effects are closely related to context effects and the two terms are sometimes used synonymously [4]. The term context effect originated in psychology and has been used to describe circumstances in which the perception of a stimulus is affected by the surrounding context, as where a gray object looks lighter against a dark background than against a light background (see, e.g., http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html). In forensic science, the term context effect has been used more broadly to describe situations in which the results of a forensic analysis are affected by the circumstances in which it is performed, and particularly by the information available to the analyst, as when an analyst becomes more likely to identify a latent print as that of a suspect when told that another analyst has already made the identification or when told that other evidence indicates the suspect made the print. The “other evidence” might be said to provide a “context” that changes the analyst’s interpretation of the data contained in the prints. Alternatively, the “other information” might be said to have changed the analyst’s expectations about the data and hence to have induced an observer effect.

The same phenomenon is sometimes described as confirmation bias. In psychology and cognitive science, the term confirmation bias refers to a human tendency to evaluate evidence in a manner that supports or confirms one’s preconceptions, such as a tendency to search out and give more weight to evidence that supports a favored hypothesis than to evidence that contradicts it [5]. In forensic science, the term confirmation bias has been used in a manner that is roughly interchangeable with the terms observer effect and context effect [6, 7].

Underlying Psychological Phenomenon

Underlying all of these terms is a basic phenomenon of human psychology – the tendency of observers to interpret data in a manner consistent with their expectations and desires. The existence of this phenomenon is well established – it has been called one of the most venerable ideas of traditional epistemology as well as one of the better demonstrated findings of twentieth century psychology [8]. An early discussion of the phenomenon can be found in the writings of Francis Bacon, who commented in 1620 that

Even earlier, Julius Caesar famously noted men’s tendency to “believe quite readily that which they wish to be true” [10]. In the twentieth century, psychologists confirmed the existence and strength of these phenomena with innumerable experiments in which people’s interpretation of a variety of stimuli, data, and forms of evidence were shown to be influenced by preconceptions and desires. The effects are greatest when the underlying data are somewhat ambiguous and when observers are influenced by strongly held expectations and motives. Extensive reviews of this literature are found in [1, 4, 5, 8, 11, 12].

The most relevant studies relate to the interpretation of scientific data. Like everyone else, scientists have a tendency to interpret data in a manner that supports theories that they favor. The tendency of
academic scientists to cling to pet theories, well past the point at which such adherence could be justified by the evidence, has been widely noted [13–15]. In his iconic book *Galileo’s Revenge: Junk Science in the Courtroom* [16], Peter Huber traced several false scientific theories to misinterpretation of data arising from uncontrolled observer effects. A scientist who is committed to a pet theory inevitably (and unconsciously) interprets data in a manner consistent with that theory. By Huber’s account, uncontrolled observer effects are one of the hallmarks of junk science (see Interpretaion: Legal Perspective).

Minimizing Observer Effects

Allowing preconceptions to influence the interpretation of data is said to be the “cardinal sin for the formal, pure scientist” [8]. Accordingly, when scientists must rely on subjective judgment to interpret the results of an experiment, they routinely take careful steps to mask or shield the person interpreting data from extraneous information that might improperly influence the interpretation. For example, scientists in most fields use “blind” or “double-blind” procedures when relying on subjective judgment to interpret data. Blind procedures are also widely used for peer-review of scientific articles, for grading of written examinations, and for other functions for which it is important to minimize observer effects.

The field of forensic science has been criticized for failing to take adequate steps to minimize observer effects [1, 4]. Through communications with police, lawyers, and other experts, forensic scientists often are exposed to information that may influence their expectations and perhaps even their hopes and desires about what a particular examination might reveal [1]. Moreover, when performing examinations forensic scientists often rely in part on subjective judgment to evaluate potentially ambiguous data [17]. Even seemingly objective procedures like DNA testing sometimes require analysts to use subjective judgment to resolve crucial ambiguities and hence could potentially be influenced by observer effects [3, 18–21]. Yet forensic scientists rarely take steps to shield themselves from extraneous information – i.e., information unnecessary for making a scientific assessment – when making comparisons or interpreting test results [1] (see also Threat Assessment: School; DNA: Degraded Samples).

The difference between forensic science and other scientific fields may stem in part from uncertainty about what types of information are necessary and relevant to a forensic science assessment. In some forensic disciplines consideration of nonscientific evidence is not only accepted but required by professional norms. For example, fire investigators are trained to consider whether a suspect had a motive for setting a fire when deciding whether to classify a fire as arson [22] (see also Fire and Explosion Investigations: Overview; Fire: Scene Investigation). There have been reports of forensic scientists relying on nonscientific evidence in other disciplines as well, such as a DNA analyst who defended a “match” that implicated the defendant in a rape case by saying: “I know I am right. They found the victim’s purse in [the defendant’s] apartment” [21]. In 1997 the US Justice Department’s Office of Inspector General reported that FBI explosives experts had relied on extraneous evidence when making key scientific determinations in a number of important cases [23]. In the first World Trade Center bombing in 1993, for example, FBI examiners had relied on the fact that the suspects had access to urea nitrate to reach the conclusion that urea nitrate had been used to make the bomb. The Inspector General condemned this reasoning as circular and biased. The FBI laboratory management agreed and pledged to take steps to ensure that the problem never happened again. Academic commentators have been unanimous in condemning the use of such nonscientific evidence as a basis for scientific conclusions [1, 6, 7, 17, 18]. They argue that the role of the forensic scientist is to offer conclusions derived from a scientific discipline, not to offer conclusions based in part on analysis of other evidence in the case (see Fire and Explosion Investigations: Overview).

Forensic scientists have become more aware of the importance of addressing observer effects as a result of a high-profile error by the FBI’s fingerprint identification unit in the Madrid train bombing case [6, 7, 24]. Three highly trained FBI latent print examiners, and a well regarded independent examiner, all identified a latent print associated with the Madrid bombing as having come from Brandon Mayfield, a resident of Portland, Oregon. The error came to light when Spanish authorities determined the latent print actually matched an Algerian suspected terrorist. The FBI acknowledged the error and apologized to Mr Mayfield. In a subsequent analysis, the agency
attributed the error in part to confirmation bias. After an initial erroneous identification, each subsequent examiner was fully aware that the latent print had previously been matched to Mayfield. This knowledge may have led them to place too much weight on similarities between the two prints while ignoring or discounting discrepancies.

**Empirical Studies of Observer Effects in Forensic Science**

Empirical studies have confirmed that observer effects can influence the interpretation of latent prints, leading to errors [25–28]. For example, in one clever study, psychologists asked five highly experienced latent print examiners to compare pairs of prints [26]. The examiners were not told that they had previously individualized these pairs (i.e., found them to match) during casework. Each examiner was instead told (incorrectly) that the prints were those that the FBI had erroneously matched in the Madrid train bombing case. Although they were instructed to “ignore all the contextual information and focus solely on the actual prints”, three of the five examiners changed their previous judgment of “match” to “no match” and a fourth changed from “match” to “cannot tell”. Only one of the five examiners consistently maintained that the prints were a “match”. In other words, after exposure to “extraneous information” suggesting that the prints should not match, four of five examiners reached a different conclusion than they had reached previously when comparing the same prints.

Studies in other forensic science domains have also found evidence of observer effects on examiners’ judgments [29]. Research on the precise mechanisms by which observer effects influence interpretation are just beginning to emerge [30]. Additional research in this important area is clearly needed.

**Case Managers and Sequential Unmasking**

There are a number of procedures for minimizing observer effects [1, 4, 31]. One method is to separate various laboratory functions and assign them to different people. A case manager who is fully informed of the facts of the case decides what to test and how to test it; a laboratory analyst who is “blind” to extraneous information analyzes and interprets the test results. To the extent possible, the analyst remains “blind” to extraneous information, such as information about the legal consequences of the judgment and nonscientific evidence in the case, when deciding whether samples “match”. Once the analyst interprets and records the test results, the case manager (who is aware of the broader facts of the case) is then responsible for placing the test results in context and assessing the compatibility of forensic observations with various theories of what occurred.

Another approach is simply to perform the analysis and evaluation of various samples in a sequence so that analysis and interpretation of early samples cannot be influenced by knowledge of the later samples. Evidentiary samples, which are generally more difficult evaluate, are analyzed first, before the analyst knows the features of the reference samples. For example, fingerprint examiners can decide whether a latent print is interpretable and which features of the latent print constitute reliable data before knowing whether those features are consistent or inconsistent with a reference print. Similarly, DNA analysts can decide which “alleles” in an evidentiary sample profile are real and which are spurious before knowing whether those alleles match up with a suspect’s. Information required to draw an ultimate conclusion is “unmasked” when needed but the analyst performs as much work as possible while “blind” to unnecessary facts [31].

**References**


4 Interpretation: Observer Effects


Related Articles

Case Assessment and Interpretation

Identification and Individualization

Interpreting Expert Opinions: History of

WILLIAM C. THOMPSON