Philosophy of Medicine
Original Research

Methods of Inference and Shaken Baby Syndrome

Nicholas Binney[1]

1 Department of Medical Ethics, Philosophy and History of Medicine, Erasmus University Medical Center, Rotterdam, Netherlands. Email: n.binney@erasmusmc.nl

Abstract

Exploring the early development of an area of medical literature can inform contemporary medical debates. Different methods of inference include *deduction, induction, abduction,* and *inference to the best explanation*. I argue that early shaken baby research is best understood as using abduction to tentatively suggest that infants with unexplained intracranial and ocular bleeding have been assaulted. However, this tentative conclusion was quickly interpreted, by some at least, as a general rule that infants with these pathological signs were certainly cases of abuse. Rather than focusing on inductive arguments, researchers today may be better off focusing on making a compelling inference to the best explanation.

1. Introduction

A fierce debate over the diagnosis of shaken baby syndrome is ongoing. Some medics are convinced that the presence of certain observations, such as unexplained intracranial and ocular bleeding, are very strong indictors that a child has been assaulted (Narang et al. 2016; Choudhary et al. 2018; Strouse 2018; Lindberg et al. 2019). Although exceptions to this rule are recognized, it is widely understood as a general rule that should influence diagnostic decisions made in medical and legal settings (Maguire et al. 2011). Others disagree, claiming that belief in the specificity of this presentation for abuse has never been established, raising concerns about misdiagnosis of abuse and miscarriages of justice (Gabaeff 2016; Lynøe et al. 2017b; Högberg et al. 2018; Findley et al. 2019). This debate has become rancorous, with frequent accusations of child abuse denialism, irresponsible scientific practice, manufacturing controversy to confuse juries, living in an echo chamber, and unwarranted etiological speculation (Strouse 2016; Choudhary et al. 2018; Duhaime and Christian 2019; Papetti, Kaneb, and Herf 2019). Skepticism about the diagnostic power of these observations has been likened to anti-vaccination propaganda (Debelle et al. 2018; Choudhary et al. 2019).

In this article, I offer an analysis of the early development of the medical literature on shaken baby syndrome, exploring the roots of this controversy and attempting to suggest a way forward. I analyze the *methods of inference* used in the early research on shaken baby syndrome. Methods of inference include *deduction, induction, abduction,* and *inference to*

ISSN 2692-3963 | Date Received 18-12-2020 | Date Accepted 21-02-2023 | Date First Published 10-05-2023
Philosophy of Medicine | DOI 10.5195/pom.2023.41 | Volume 4 | Issue 1 | pp.1–33

the best explanation. Induction, as I describe it here, is an inference based on the frequency with which events are correlated. Abduction and inference to best explanation, by contrast, are inferences based on the ability to explain surprising observations (Douven 2022).[1] Both frequency-based inferences and explanation-based inferences can be very strong arguments. Alternatively, they may only provide weak support to their conclusions. They also differ in the role they play in science. Induction and inference to the best explanation are used to try and establish that a conclusion is true. Abduction, however, is often characterized as a way of offering possible conclusions for further investigation. A perfectly good abduction is not necessarily conclusive. I argue that the arguments made in the early shaken baby syndrome literature are best understood as abduction in this sense, rather than induction or inference to the best explanation. As such, these arguments should not have been interpreted as supporting their conclusions so strongly as to put them beyond all reasonable doubt, but they were quickly interpreted in this way. I suggest that this interpretation of the early literature may have set off a self-reinforcing cycle, in which circular arguments made the general rule described above make it seem much more securely established than it is. Given this, clarifying the methods of inference used in medical arguments is an important task for the philosopher of medicine.

I begin by introducing a debate within the medical literature on the diagnosis of shaken baby syndrome (section 2) and discussing what these researchers observed and claimed (section 3). Many doctors hold that a strong statistical correlation between intracranial and ocular bleeding and abuse has been established by induction. I argue that, at least initially, this was not the case (section 4). The early shaken baby syndrome researchers did not observe that children with certain signs have been violently assaulted, thereby establishing by induction that children with these signs have been assaulted. Rather, they observed that there are children who have been shaken who had these certain signs, suggesting by abduction that children with these signs may have been assaulted (section 5). I also show how these early researchers used abduction as a sense-making inference to the same effect (section 6). I describe inference to the best explanation as another explanation-based inference used to justify conclusions (section 7). I suggest that these arguments should be understood as abductions, which support their conclusions tentatively. Other researchers, such as Stephen Ludwig and Matt Warman (1984) and Ann-Christine Duhaime et al. (1987), quickly embraced the belief that patients with these signs have almost always been violently assaulted. I suggest that they did so because they interpreted the work of early researchers as making either a compelling induction or an equally compelling inference to the best explanation, rather than an abduction (section 8). How doctors respond to new evidence will depend on how they understand the arguments that support their beliefs. Paying close attention to the development of beliefs used to diagnose shaken baby syndrome can help us understand why people believe the things they do, and thereby inform contemporary debates about this condition. I conclude that it may be more profitable for researchers to focus on making compelling inferences to the best explanation, rather than continuing to focus on induction.

[1] Some philosophers refer to all explanation-based inferences as abduction (Niiniluoto 2018; Williamson 2018; Douven 2022). It is useful for me to distinguish between explanation-based inferences that tentatively suggest new hypotheses (abduction) and those that try to establish claims as true (inference to the best explanation). Also note that some philosophers argue that abduction does not need to be explanatory (Magnani 2009, 63), but I do not take this position here.

2. Shaken Baby Syndrome

For many years, a group of medics has championed the view that the great majority of infants who present with subdural hemorrhage, retinal hemorrhage, and encephalopathy (brain damage) have been violently assaulted, especially if there is no history of serious trauma to the head. These patients are said to be victims of abusive head trauma, or non-accidental head injury. This collection of pathologies was at one point referred to as "shaken baby syndrome":

> The act of shaking leading to shaken baby syndrome is so violent that individuals observing it would recognize it as dangerous and likely to kill the child. Shaken baby syndrome injuries are the result of violent trauma. The constellation of these injuries does not occur with short falls, seizures, or as a consequence of vaccination. (American Academy of Pediatrics: Committee on Child Abuse and Neglect 2001, 206)

> Subdural hemorrhage or retinal hemorrhage can be found in infants in situations other than abuse; in the great majority of cases, however, the combination of subdural hemorrhage and retinal hemorrhage is a result of SBS [shaken baby syndrome]. (Matschke et al. 2009, 216)

> The findings of subdural hematoma, retinal hemorrhage and hypoxic–ischemic encephalopathy remain highly suggestive of shaken baby syndrome, particularly in the absence of evidence of an impact injury. (Strouse 2018, 1043)

For the sake of simplicity, I refer to the findings of subdural and retinal hemorrhage along with encephalopathy, without any signs of external head injury and without a history of a serious trauma to the head, as *unexplained intracranial and ocular bleeding*. This group of researchers claim as an observed fact that infants with unexplained intracranial and ocular bleeding have almost always been violently assaulted. A recent consensus statement on the diagnosis of abusive head trauma cited studies which found that "any combination of three or more of the significant diagnostic features [which include subdural bleeding, retinal bleeding, and no history of trauma] yielded a positive predictive value of 85%" (Choudhary et al. 2018, 1052). Others have claimed that, in patients with intracranial and retinal bleeding, having "no history of trauma had a specificity of 100% and positive predictive value of 100% for AHT [abusive head trauma]" (Narang 2011, 560). Although they concede that there are a few alternative diagnoses that always need to be considered, these researchers claim to have *observed* that unexplained intracranial and ocular bleeding is virtually pathognomonic for violent assault once these few alternative diagnoses have been taken into account. Consequently, the diagnosis of abuse functions as a default: "The question to be answered is, 'Is there a medical cause to explain all the findings or did this child suffer from inflicted injury?'" (Choudhary et al. 2018, 1059).

But is this true? Of course, if researchers assume that patients with unexplained intracranial and ocular bleeding have just about always been abused, almost no patients with this presentation will be seen as *not* abused. If the diagnosis of abuse is made on this basis, unexplained intracranial and ocular bleeding will appear perfectly specific for abuse,

even though it may not be. Such studies are uninformative about whether unexplained intracranial and ocular bleeding is actually specific for abuse. Indeed, the main criticism of much of the research into shaken baby syndrome/abusive head trauma is that is makes circular arguments, in which the specificity of unexplained intracranial and ocular bleeding is not observed, but rather is assumed from the outset (Donohoe 2003; Findley et al. 2012; Lynøe et al. 2017b).[2]

Sandeep Narang (2011, 561) has made an interesting counterargument to the charge of circular reasoning: Circular arguments do not generate new beliefs; they reinforce a belief already held. Sometime in the past, medics did not believe that the great majority of patients with subdural and retinal hemorrhage had been assaulted. And yet, following the arguments made in early studies on the subject, many came to believe that this was true. Hence, these early researchers cannot have made circular arguments. How was the link between unexplained intracranial and ocular bleeding and violent assault made in the first place?

This is a very good question, which can be addressed by looking at this early literature to see how the claim that the great majority of patients with unexplained intracranial and ocular bleeding had been assaulted was supported by empirical observation.

3. Observations and Claims Made by Early Shaken Baby Syndrome Researchers

The key papers in this early literature are by John Caffey (1972, 1974), A. Norman Guthkelch (1971), and C. Henry Kempe and colleagues (1962) (Narang 2011; Choudhary et al. 2018). It is important to consider the patients these researchers looked at, what they observed of these patients, and what they concluded from these observations. Then we can think about how their conclusions were supported by their observations.

All of these papers present cases of infants who had been violently assaulted. Kempe and colleagues (1962) identified several hundred instances of what they called the "battered child syndrome," using bone fractures, other injuries and subdural hemorrhage. Gulthkelch (1971) presented a smaller series of 23 cases of "proved or strongly suspected assault," including both children who were battered and children who were shaken. Caffey gathered accounts of 27 cases of violent assault in small children, who had subsequently become ill or died, which he described as "several convincing recorded examples of pathogenic and even fatal shaking" (Caffey 1972, 163). He also re-described six infants with bone fractures and subdural hemorrhage that he had identified in 1946 as victims of violent assault by shaking and discussed the pathological signs of many other abused infants found in medical literature (Caffey 1972, 1974). He drew particular attention to a group of infants who became ill after having been looked after by the same nanny, who later admitted to shaking them.

> By far the most extensive anecdotal proof of pathogenic manual WLS [whiplash shaking] comes from the confessions to the savage shaking of dozens of infants by an infant-nurse who whiplashed three infants to death, maimed two others, and shook uncounted others during a period of nine years. (Caffey 1974, 397)

[2] Donohoe (2003), Findley et al. (2012), and Lynøe et al. (2017b) have drawn substantial criticism (Hellgren et al. 2017; Lucas et al. 2017; Lynøe et al. 2017a, 2018; Saunders et al. 2017; Laurent-Vannier et al. 2018; Cohen 2019; Lynøe and Eriksson 2019; Janson 2020). Nevertheless, the shaken baby syndrome literature is riddled with circular arguments.

So, each of these authors present cases of infants whom they were sure were victims of a violent assault. Each observed that subdural hemorrhage occurs in assaulted patients. They also note that many assaulted infants developed this intracranial bleeding without other injuries to the skull or other parts of the body. According to Kempe and colleagues, "subdural hematoma, with or without fracture of the skull, is, in our experience, an extremely frequent finding even in the absence of fractures of the long bones" (1962, 18). Guthkelch found 13 cases of subdural hemorrhage, in 5 of which "no evidence of application of direct violence to the head was forthcoming" (1971, 431). Caffey also noted that many of these cases had retinal hemorrhages as well as subdural hemorrhages (Caffey 1972, 167; 1974, 399–400). He emphasized that subdural and retinal hemorrhages could occur together in shaken infants in whom there was no evidence of injury to the outside of the head or to the neck:

> The most characteristic pattern of physical findings in the whiplashed infant is the absence of external signs of trauma to the head and the soft tissues of the face and neck, and of the facial bones and calvaria, in the presence of massive traumatic intracranial and intraocular bleedings. (Caffey 1974, 399)

Additionally, Kempe and colleagues (1962) document cases that they are sure are cases of abuse, which have subdural bleeding, fractures, and *no history of trauma*. Caffey did the same: "Usually there is no history of trauma of any kind" (Caffey 1974, 402; see also Caffey 1972, 165).

It is important to note that these researchers present almost no observations of patients selected because they have subdural and retinal hemorrhage of any cause, to see whether these arise without trauma and without violent assault. Guthkelch (1971) compares two series of patients with head trauma to another series of battered patients, to show that subdural bleeding, while present in all three series, occurs less frequently in the head trauma patients than in the battered patients. Kempe and colleagues (1962) focus on battered infants. Caffey mentions idiopathic retinal hemorrhage of the newborn infant, only to say that it is common, usually resolves within one week, and is not due to birth trauma (1974, 400). Caffey (1972, 166; 1974, 400) also mentions a case series of subdural hemorrhage in infants, in which 54% of patients had no history of trauma at all (Ingraham and Matson 1954). However, as we shall see, he does not interpret this as the observation of patients with subdural hemorrhage who have not been abused. These early shaken baby researchers do not try to observe whether patients with intracranial and ocular bleeding have been violently assaulted.

So, what these researchers observed is that *there are patients who have been violently assaulted who have unexplained intracranial and ocular bleeding*. The main conclusions of these papers are stated clearly. Kempe and colleagues advised that the battered child syndrome

> should be considered in any child exhibiting evidence of fracture of any bone, subdural hematoma, failure to thrive, soft tissue swelling or skin bruising, in any child that dies suddenly, or where the degree and type of injury is at variance with the history given regarding the occurrence of the trauma. (1962, 17)

Guthkelch reached a similar conclusion: "One must keep in mind the possibility of assault in considering any case of infantile subdural hematoma, even when there are only trivial bruises or indeed no marks of injury at all, and inquire, however guardedly or tactfully, whether perhaps the baby's head could have been shaken" (1971, 431). So, these researchers concluded that should clinicians see any of these signs, even without signs of external injury to the head or a history of serious trauma, they should suspect abuse.

Caffey's main claim was that shaking small children could produce subdural and retinal hemorrhages, in addition to long bone fractures (1972, 1974). His main objective was to convince parents that shaking children is dangerous. This plea was combined with the conclusion that shaking is a potential explanation for subdural and retinal hemorrhage in infants: "WLS [whiplash shaking] of the head was the reasonable explanation for the presence of bilateral subdural hematomas and bilateral intraocular hemorrhages, combined with the concurrent absence of external signs of trauma to the head and neck and to the extremities in more than half of our early cases" (Caffey 1974, 396–397).

The main conclusion of these studies can be summarized as follows: *unexplained intracranial and ocular bleeding indicates violent assault.*

4. Early Shaken Baby Researchers Did Not Use Induction

This conclusion is similar to the claims made by contemporary researchers on shaken babies. But how was it arrived at? Perhaps the early researchers observed that unexplained intracranial and ocular bleeding is associated with violent assault and associated in such a way that patients with these signs just about always have this disease. To do this, the early researchers would have to have observed patients with unexplained intracranial and ocular bleeding and seen that they had just about always been assaulted. As discussed, the early researchers did observe patients with these signs who had been assaulted. So, perhaps, we might say that they were making this sort of argument. But they were not. To grasp this, we must think a little more carefully about different sorts of argument, and especially about the methods of inference used to make them.

Methods of inference are the means by which the premises inform the conclusion of an argument. Premises are observations made by researchers, or other claims that the researchers take for granted. The sort of argument we are considering just now is an *induction,* which has several varieties.[3] One form of induction is *universal generalization,* which takes particular observations of something being the case and transforms them into a general rule that this thing will always be the case (Robinson 2021). This general rule is the conclusion of the argument. Induction is a *more-of-the-same* inference: it assumes that what has been observed in particular cases will apply generally. So, the observation of one ripe banana that is yellow, and another, and another, and so on, leads to the general rule that all ripe bananas are yellow. Similarly, the observation of one patient with a certain set

[3] Induction is sometimes used to refer to any method of inference that is not deduction, and thus to include abduction and inference to the best explanation under one umbrella (Flach and Kakas 2000). Here, I draw a distinction between induction, abduction, and inference to the best explanation. See Henderson (2020), Climenhaga (forthcoming), and Johannesson (2022) for more thorough discussions of different sorts of induction. I am using the term to refer to universal and statistical generalization here.

of signs has a particular disease, and another, and so on, leads to the general rule that all patients with these signs have this disease.

Induction is often contrasted with deduction. The conclusion of a deductive argument must be true if the premises of the argument are true. In medicine, doctors may use deductive arguments to diagnose a disease from its diagnostic criteria. Given that all patients with certain signs have a particular disease, if a particular patient has those signs, that patient must have that disease. In contrast to this, the conclusion of an inductive argument need not be true, even if the premises are. So, even though it may be true that all the swans I have ever seen are white, this does not mean that all swans must be white. Even though all the patients I have ever seen with certain signs have a particular disease, this does not mean that all patients with those signs must have that disease. Even so, the more I observe patients with those signs that have that disease, the more confident I become that the general rule that all patients with those signs have that disease is true.

Induction does not have to lead to general rules that are perfect. Another form of induction is *statistical generalization* (Robinson 2021). We might observe that only half of patients with a particular disease have a certain set of signs. Given this, we might conclude that, in general, half of patients with this disease have a certain set of signs. The shaken baby syndrome researchers above claim to have made particular observations that between 85% and 100% of infants with unexplained intracranial and ocular bleeding have been violently assaulted, leading to their general conclusion that between 85% and 100% of patients with these signs have been violently assaulted. Induction turns particular observations of something being the case into a general rule that this will always be the case.

Inductions are certainly not infallible. With a more-of-the-same inference, to be confident that the results of an old induction will apply to some new situation, we must be confident that the new situation is the same as the old one. This requires background knowledge, judgment, and interpretation about relevant differences. [4] Diagnostic researchers regularly find that the sensitivity and specificity of tests vary unexpectedly in clinical populations that were once thought of as the same. However, my purpose here is not to question the inductions researchers may have made, but rather to question whether an induction has been made at all.

When making inductions, we must be very careful to note the group of things that we are studying. When we observe that ripe bananas are yellow, we are studying *ripe bananas,* and not the group of *yellow things*. Studying ripe bananas, and observing that they are all yellow things, allows us to induce that all ripe bananas are yellow. These observations do not, however, tell us about how many yellow things are bananas. Even if all ripe bananas are yellow, there might be many yellow things that are not bananas, such as lemons. Alternatively, there may be almost no yellow things that are not bananas. If we only study bananas, we cannot tell.

The same goes for the study of diseases and their signs. Figure 1 uses a blue circle to represent patients with a particular disease, and a green circle to represent patients with a certain set of signs. The red area represents the patients who have both the disease and those signs. If we study the blue circle, we can observe the percentage of that circle which is red. If we study the patients with the disease, we can observe that a certain percentage of

[4] All observations involve interpretation in this sense (Hanson 1958). All observation is theory laden. Indeed, some have argued that inductions are all actually inferences to the best explanation, because of explanatory assumptions hidden in all observations (Harman 1965).

them have the signs. However, if we only study the blue circle, we cannot observe the percentage of the green circle that is red. If we only study the patients with the disease, we cannot observe the percentage of patients with those signs that have the disease. There may be many patients with those signs who do not have the disease, there may be none; we cannot tell. Even though we have observed patients with those signs who have the disease (the red area), we cannot make inductive inferences about the percentage of patients with those signs that have the disease (the percentage of the green circle that is red), because we have not studied the patients with those signs (the green circle).

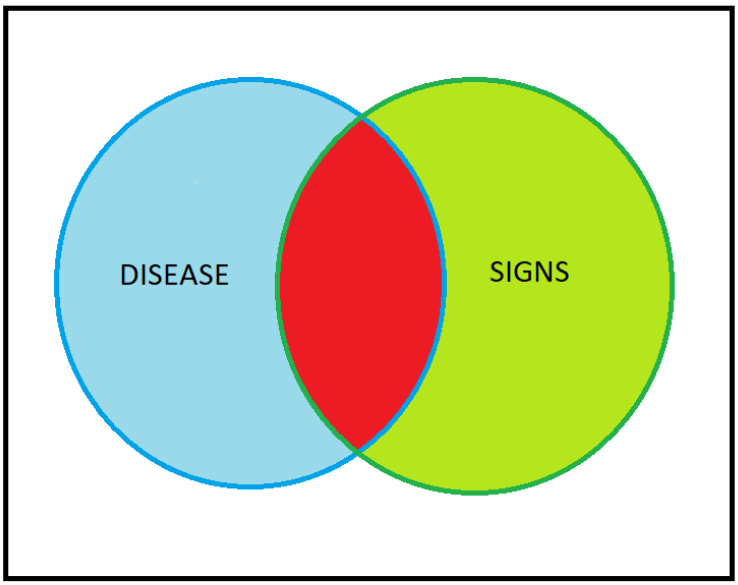

Figure 1. Overlapping sets of patients with a disease and patients with certain signs.

So, what did the early shaken baby researchers study? As discussed, both Kempe et al. (1962) and Guthkelch (1971) studied patients who had been violently assaulted and observed how many of them had unexplained intracranial and ocular bleeding. They studied patients with a particular disease (the blue circle) and observed how many of them had a certain set of signs (the red area). So, Guthkelch (1971) presented a case series of 23 assaulted infants and found that 13 of them had subdural bleeding. From these particular observations, we might induce the general rule that 58% of assaulted infants will have subdural bleeding. Similarly, Kempe et al. (1962) observed that subdural hemorrhage was a frequent finding among victims of assault. However, we cannot induce a general rule about the proportion of infants with subdural hemorrhage who have been assaulted from these observations, because this work did not study patients with subdural hemorrhage. These early researchers only studied patients with a particular disease, without studying patients with a certain set of signs.

Caffey (1972, 1974) was even more selective. Caffey presented a case series of patients who had been assaulted and then became ill or died. He did not study the group of patients with a particular disease as a whole (the blue circle), or the group of patients with a certain set of signs as a whole (the green circle). Caffey only studied patients with a particular

disease if they also had these signs (he studied the red area only). Consequently, Caffey's observations cannot be used to induce general conclusions about either infants who have been violently assaulted or patients with unexplained intracranial bleeding.

And yet, as discussed, these early researchers do seem to draw conclusions about something that they have not studied. They (at best) made observations about whether or not patients with a particular disease have certain signs, but drew a conclusion about whether or not patients with certain signs have a disease. If these early researchers claimed to have observed that unexplained intracranial and ocular bleeding is so tightly associated with violent assault that patients with these signs have almost always been assaulted, it seems as though they were mistaken. However, they did not make such claims. They made no such induction. They made a completely different sort of argument, called an abduction.

5. Early Shaken Baby Researchers Used Abduction

Abduction may be a less well-known method of inference than induction, but it is not less important. It is also widely used in medicine. Doctors should recognize this form of inference from many areas of their own work (Upshur 1997; Magnani 2004; Bolton 2015; Chiffi 2020, 54–59; Karlsen, Hillestad, and Dysvik 2021). For example, should doctors see a patient with a nocturnal cough and shortness of breath when walking upstairs, they might think to themselves, "Patients in heart failure sometimes have a nocturnal cough and get short of breath walking upstairs, so perhaps this patient has heart failure." Should medics see a patient who is urinating much more than normal, and who has sugar in their urine, they might think to themselves, "Patients with diabetes sometimes urinate more than normal and have sugar in their urine, so perhaps this patient has diabetes." Such doctors are not arguing that patients with these symptoms inevitably have those diseases. Their argument is not a deduction. Neither are they arguing that the last hundred times they saw a patient present like that, they had they had diabetes, so that is likely to be the case again, although such arguments could be made. Their argument is not an induction. They are arguing that they have seen cases with those diseases present like this before, and this gives them reason to believe this could be happening again. This sort of argument is an abduction.

The philosophical study of this form of inference is indebted to the American philosopher Charles Sanders Peirce, who worked in the nineteenth and early twentieth centuries. Peirce introduced the discussion of this form of inference to modern philosophy, which he worked on for more than 50 years. Over this time, Peirce's views on abduction developed, and he described abduction in several different ways (Fann 1970; Flach and Kakas 2000; Paavola 2005; Plutynski 2011; Pietarinen and Bellucci 2014). Initially, he referred to it as the method of hypothesis, and then as abduction, and then as retroduction. I will use the term "abduction" here to refer to any of these. Some of Peirce's formulations of abduction are valuable here. For example, he discussed an inference one might make about where a handful of white beans on a table might have come from:

> Suppose I enter a room and there find a number of bags, containing different kinds of beans. On the table there is a handful of white beans; and, after some searching, I find one of the bags contains white beans only. I at once infer as a probability, or as a fair

> guess, that this handful was taken out of that bag. This sort of inference is called making an hypothesis. It is the inference of a case from a rule and a result (Peirce, *CP* 2.623).[5]

As the bag only contains white beans (which is what Peirce calls the "rule"), in the event that a handful of beans is drawn from this bag (the "case"), they would all inevitably be white (the "result"). As there is a handful of white beans on the table (the result), and we know that the bag only contains white beans (the rule), it is reasonable to suggest (it is "a fair guess") that the beans are from that bag (the case). Peirce represents this inference as follows:

> HYPOTHESIS.
> Rule.—All the beans from this bag are white.
> Result.—These beans are white.
> ∴ [therefore]
> Case.—These beans are from this bag. (Peirce, *CP* 2.623)

Applying this to diagnostics, if we see patients with a certain set of symptoms (the result), and if all patients with a disease have that set of symptoms (the rule), it is reasonable to suggest (it is "a fair guess") that these patients have this disease (the case).[6] This might be a good way to represent the reasoning process by which physicians come to suspect that a patient has a disease as they compile a list of differential diagnoses.[7]

However, the rule that all patients with the disease need to display these symptoms is perhaps too restrictive. Peirce did indeed seem, initially at least, to think that abductions (hypotheses) are only reasonable if the result follows deductively or "syllogistically" from the rule and the case (Pietarinen and Bellucci 2014, 356). This requires that all the beans from the bag be white; otherwise, a handful of beans drawn from the bag would not inevitably be white. If only some beans from the bag are white, why would anyone expect that a handful of beans from this bag should all be white? Why would someone having taken a handful of beans from this bag be a good explanation for all the beans on the table being white?

Even so, a group of patients with a certain set of symptoms need not be thought of as a random sample (a handful) of patients with a particular disease. After all, this group of patients might be assembled precisely because they have those symptoms. There is no reason to suspect that a particular group of patients with a disease should have the same frequency of symptoms as the group of all patients with that disease. In order to suspect that the patients might have a disease, it is enough to know that at least some patients with a disease have those symptoms. Even if only some patients with the disease have those symptoms, it may be fair to guess that patients with those symptoms have that disease.

Peirce provides other formulations of abduction that seem not to require such a strict rule. According to this scheme, so long as a kind of thing (perhaps a disease "M") has a

[5] *CP* 2.623 refers to *Collected Papers of Charles Sanders Peirce,* volume 2, paragraph 623.
[6] This sense of abduction has been referred to as Peirce's syllogistic theory, which can be seen as a precursor to the current use of abduction in logic programming and artificial intelligence (Flach and Kakas 2000, 5).
[7] Some distinguish between *selective* abduction, which compiles a list of plausible hypotheses from an already existing catalogue, and *creative* abduction, which proposes entirely new hypotheses (Magnani 2004). Both processes are important to many areas of medical practice. For more discussion of different types of abduction, see Park (2015). For a contemporary attempt to capture the logical scheme of abduction, see Woods (2013, 379).

certain set of characteristics (perhaps the symptoms "P[1], P[2], P[3], etc."), it is a fair guess that an object that has these characteristics (patient "S") is that kind of thing (has the disease "M"):

> A well-recognized kind of object, M, has for its ordinary predicates P[1], P[2], P[3], etc., indistinctly recognized.
> The suggesting object, S, has these same predicates, P[1], P[2], P[3], etc.
> Hence, S is of the kind M. (Peirce, *CP* 8.64)

Even though the conclusion of this argument is "S is of the kind M," it is important to remember that for Peirce abduction only offered a fair guess, not conclusive evidence, about what is the case. That "S is a kind of M" is suggested as a hypothesis, and not put forward as a claim about what is true. Furthermore, although Peirce maintained that this "is the hypothetic inference in form" (*CP* 8.64), he was not suggesting that anyone actually expressed their thinking like this. Peirce held that this inference can be carried out without much deliberation, almost unconsciously, for both scientific and everyday purposes. I take this as one useful way of representing abduction here.

Medics might know that at least some patients with a particular disease have a certain set of signs. Given this, medics can infer that patients with those signs might well have that disease. We all know that ripe bananas are yellow things; therefore, yellow things might be ripe bananas. These arguments are abductions in the above sense.

Like induction, abduction goes beyond its premises. There are yellow things that are not bananas, such as lemons. Hence, a yellow thing might not be a banana at all, as it could be a lemon. There are patients with a certain set of signs who do not have the disease. Hence, patients with those signs might have a different disease that presents in a similar way. Nevertheless, abduction and induction are different. Abductions allow inferences about groups of things that have not been directly observed to be made, while inductions are limited to generalizations about what has been directly observed. Abductions allow us to infer that a yellow thing could be a banana, even though we have not studied the whole group of yellow things, and that patients with certain signs might have the disease even though we have not studied the whole group of patients with those signs. "But the essence of an induction is that it infers from one set of facts another set of similar facts, whereas hypothesis infers from facts of one kind to facts of another" (Peirce, *CP* 2.642).

The early shaken baby researchers argued by abduction. Kempe and colleagues (1962) used their observations of battered infants to conclude that when unexplained intracranial bleeding is observed, violent assault should be *considered*. Guthkelch (1971) used his observations of assaulted children to conclude that infants with unexplained intracranial bleeding had *perhaps* been assaulted. These researchers observe that at least some patients who have been violently assaulted have unexplained intracranial and ocular bleeding and infer from this that patients with unexplained intracranial and ocular bleeding might have been assaulted. This is an abduction. Caffey also argued in this way, but his conclusion was not so tentative. Perhaps Caffey thought that this conclusion should not just be *considered,* but rather should be *accepted as true*. Even so, Caffey conceded that the evidence he used to suggest the pathogenicity of shaking was incomplete and thus "did not lend itself to satisfactory statistical analysis" (Caffey 1972, 168–169). His was not a frequency-based argument. Rather, Caffey (1972, 1974) argued that shaking *was the reasonable explanation*

for how an infant could have intracranial bleeding without signs of external injury to the head, and without a history of serious trauma to the head. Before discussing how explanation-based inferences might justify beliefs, another way of presenting the form of an abduction should be discussed.

6. Abduction as a Sense-Making Inference

The schemes above present abduction as a way to infer what is the case (diagnosis) from a rule (a disease produces these symptoms) and a result (the patient's symptoms). While this is true, abduction is also more than this.

Abduction can also be understood as an inference that makes sense of a confusing set of observations. It is the sort of thing that a detective might do at a crime scene. They see the empty jewellery box, they see the muddy footprints on the carpet, they see the broken window, they see the footprints in the flower bed outside, and they see the butler's muddy shoes, and, from all this put together, they suspect that the butler did it. They take a disparate and confusing collection of observations, which serve as premises, and unite them into an intelligible and coherent whole, thereby drawing their conclusion. That the confusing observations can be made intelligible by accepting that the butler did it gives the detective reason to suspect that this is what happened.

As his thinking about abduction matured, Peirce came to describe abduction as a way of making sense of a confusing set of observations by inventing entirely new explanations and classifications. Instead of starting with a well-known rule, or a well-recognized object with well-known characteristics, Peirce described a process that started with confusing and chaotic experiences that people try to explain, thereby making sense of the chaos and producing order in what was once disordered:

> A mass of facts is before us. We go through them. We examine them. We find them a confused snarl, an impenetrable jungle. We are unable to hold them in our minds. We endeavor to set them down upon paper; but they seem to be so multiplex intricate that we can neither satisfy ourselves that what we have set down represents the facts, nor can we get any clear idea of what it is that we have set down. But suddenly, while we are poring over our digest of the facts and are endeavoring to set them into order, it occurs to us that if we were to assume something to be true that we do not know to be true, these facts would arrange themselves luminously. That is abduction. (Peirce 1992, 531)[8]

The ability to explain this jumble of disordered experiences is what Peirce thought licensed the suggestion that this new way of looking at things is true. Peirce tried to capture this in yet another scheme for abduction:

> The surprising fact, C, is observed;
> But if A were true, C would be a matter of course,
> Hence, there is reason to suspect that A is true. (Peirce, *CP* 5.189)

[8] This is a passage from Peirce's seventh lecture on pragmatism, "Pragmatism as the Logic of Abduction," delivered at Harvard University in 1903. Reproduced in Peirce (1992, 531). This passage is from a deleted section of that lecture.

Again, it is worth emphasizing that, for Peirce, this explanatory power only provides good reason to suspect that something is true. Peirce did not hold that abduction could license firm belief in the suspected explanation. Abduction could generate new explanations, and could provide reasons for investigating these explanations further, but he denied that abduction itself could provide grounds for accepting the explanation as true. Peirce insisted that abduction was a weak form of inference, certainly in comparison to induction: "Induction is, plainly, a much stronger kind of inference than hypothesis; and this is the first reason for distinguishing between them" (Peirce, *CP* 8.65). This is not to say that this guess is just a random stab in the dark. Rather, Peirce argued that people were quite good at guessing what was the case, as is evidenced by human beings being far more successful at activities like science than we would be if we were just randomly guessing at the truth (Paavola 2005). Even so, he maintained that all abduction provided was a fair guess at what was true: "Deduction proves that something must be; Induction shows that something actually is operative; Abduction merely suggests that something may be" (Peirce, *CP* 5.171).

This does not mean that Peirce thought that abduction was useless. On the contrary, he thought it was essential to science. He came to believe that science was a process that involved abduction, deduction, and induction in sequence: Abduction proposes hypotheses, deduction works out the consequences of these hypotheses, and induction looks to see if these consequences occurred (Fann 1970).[9] While abduction and induction are both ampliative, in that they draw conclusions that go beyond their premises, abduction is special because it can propose entirely new kinds of object to feature in explanations. Induction, by comparison, is derivative; it only checks for correlations between objects already proposed. Deduction only restates what is already contained in the premises. Thus, Peirce came to think of abduction as the least secure but the most fruitful stage of scientific inference (Pietarinen and Bellucci 2014, 354; Peirce, *CP* 8.387). "All the ideas of science come to it by the way of Abduction. Abduction consists in studying facts and devising a theory to explain them. Its only justification is that if we are ever to understand things at all, it must be in that way" (Peirce, *CP* 5.145).

Peirce also held, at least at one point, that abduction is valuable because sometimes we cannot carry out inductions, and thus have no choice but to make abductions if we are to make inferences at all. "Hypothetic reasoning infers very frequently a fact not capable of direct observation. It is an hypothesis that Napoleon Bonaparte once existed. How is that hypothesis ever to be replaced by an induction?" (Peirce, *CP* 2.642).

These ideas are applicable to the case of shaken baby syndrome. Unexplained intracranial and ocular bleeding is confusing. To resolve this confusion, new categories of patient, such as shaken baby syndrome, and new explanations, such as shaking, need to be produced. Abduction is capable of doing this. Furthermore, as the alleged instances of shaking happened in the past and cannot be observed directly, researchers may be forced to use abduction to infer what is happening in these confusing cases.

The early shaken baby researchers also used such a sense-making inference. Both Guthkelch and Caffey note that "subdural haematoma is one of the commonest features of the battered child syndrome, yet by no means all the patients so affected have external marks of injury on the head" (Guthkelch 1971, 430). The "yet by no means" here indicates that this finding is somewhat confusing. How can an infant with no signs of injury to their

[9] Models for medical practice have been built on this Peircian foundation (Ramoni et al. 1992).

skull suffer with bleeding inside their head? Caffey calls "the presence of massive traumatic intracranial and intraocular bleedings" without any external injuries to the head an "extraordinary diagnostic contradiction" (1974, 399). So, these researchers identify a confusing set of clinical observations in need of explanation.

Guthkelch (1971, 430) and Caffey (1974, 401–402) explain these confusing observations by drawing on research from Ayub K. Ommaya and colleagues (1968, 1969), who simulated road traffic accidents using primate subjects. They put these primate subjects into a seat that could slide on rails and slammed a heavy sled into the back of it to simulate a rear-impact crash. This whiplashed the subjects' heads backward and forward and Ommaya and colleagues observed that this whiplashing could produce subdural hemorrhage. They also noted two cases of real road traffic accidents that had produced subdural hemorrhage by a similar mechanism (Guthkelch 1971, 430). The whiplashing movement caused the brain to move relative to the skull, tearing the veins that ran between the two, without any direct impact to the outside of the head. Gulthkelch argued that the subdural bleeding could be explained on this basis: "It is now submitted that the conditions which are known to exist in many cases of the battered child syndrome are particularly favourable to the production of subdural haematoma in infants by an essentially similar mechanism" (1971, 430).

This argument starts with a confusing set of observations and uses an experimental model of road traffic accidents to propose an explanation for these confusing observations. That this explanation can make sense of these confusing observations is used to support the view that this is indeed the correct explanation of these confusing findings. This argument is also an abduction, but instead of appealing to cases, this sort of abduction appeals to the ability to make sense of a confusing set of observations to support the inference made.

Both Caffey (1972, 1974) and Guthkelch (1971) also found they could explain other confusing observations in this way. Fractures to the humerus could be explained by the parent gripping the child by the upper arms in order to shake it. Many subdural hemorrhages in these infants were bilateral, occurring over both left and right sides of the brain (Guthkelch 1971; Caffey 1974, 400). When held by the upper arms, the infant would be shaken backward and forward, producing a symmetrical backward and forward motion of the infant's head, which would affect both sides of the head equally, producing a bilateral subdural hemorrhage. The bilateral subdural hemorrhages *made sense* if the infants had been shaken. Guthkelch also noted that, in case series presented by other authors, subdural hemorrhage was more frequent in case series of assaulted infants than it was in case series of infants with other forms of head injury: "Such an hypothesis might also explain the remarkable frequency of the finding of subdural hemorrhage in battered children as compared with its incidence in head injuries of other origin, and the fact that it is so often bilateral" (Guthkelch 1971, 430). This is an abduction.

Caffey drew attention to the frequent finding of retinal lesions found in supposedly battered children, and how other researchers felt these could not be explained by battering: "They propose that some of the affected infants are the victims of over-vigorous *manipulations,* not battering. We agree with them and believe that many of these infants are whiplash-shaken rather than beaten, especially those with intracranial and intraocular bleedings" (Caffey 1974, 399). Caffey argued that the combination of subdural and retinal hemorrhage was better explained by shaking than by battering. Caffey also noted the finding by Franc D. Ingraham and Donald D. Matson (1954) that 54% of patients with subdural hemorrhage had no history of trauma, which was especially confusing because subdural

hemorrhage was believed to "always be of traumatic origin" (Caffey 1972, 166). This finding would make sense if subdural hemorrhage was produced by shaking that the parents then did not disclose. "The absence of a history of trauma of any kind is significant and suggests that whiplash shaking may be the cause in many patients" (Caffey 1974, 400). Caffey then concluded: "Many of these facts are better explained theoretically on the basis of repeated subdural bleedings induced by repeated whiplash shaking which causes progressive cumulative changes in the hematomas over several weeks or months" (1974, 401). All these disparate and confusing facts could be made sense of as a coherent whole if these infants were shaken, which gave these researchers reason to believe that these infants were in fact shaken. This argument could be an abduction, understood as a sense-making inference, rather than as an inference from previous cases. However, it could also be another kind of explanation-based inference, known as inference to the best explanation (IBE).

7. Inference to the Best Explanation

If we follow Peirce, explanatory inferences like this should be seen as weak methods of inference, which are little better than a guess. While this may often be true, it need not be the case. Consider the use of DNA evidence in murder trials. Say the defendant's DNA was found on the murder victim's body. This observation could provide compelling evidence to support the conclusion that the defendant is the murderer. But why? This is because the presence of the defendant's DNA on the victim's body would make sense if the defendant was indeed the murderer. If alternative explanations (for example, the DNA was deposited when the defendant and the victim met earlier in the day, or the defendant is being framed by the real murderer) are considered poor explanations, this could be a very strong form of inference indeed. So strong that, in some countries, the death penalty is applied or revoked based upon explanation-based inferences like this.

Philosophers have argued that some of the most important inferences made in the history of science have been made by making sense of previously confusing observations. The discovery of the elliptical orbit of Mars, the oxygen theory of combustion, and the theory of evolution by natural selection were all driven by such considerations (Hanson 1958; Thagard 1978; O'Rorke, Morris, and Schulenburg 1990; Okasha 2000; Douven 2022). Johannes Kepler found he could make sense of the confusing mass of planetary observations supplied by Tycho Brahe, if he assumed that Mars moved in an ellipse. Antoine Lavoisier found that he could make sense of confusing observations that some substances became heavier as they burned, if he assumed that they combined with a substance in the air. Charles Darwin found he could make sense of confusing observations concerning the distribution, anatomy, embryology, fossil record, and behavior of different species by assuming that they evolved from a common ancestor in response to selective pressure from their environment. Furthermore, philosophers argue that belief in these canonical scientific theories was not only discovered by such sense-making inferences, but also justified by them. Reflecting on criticism of his argument, Darwin commented in the sixth edition of *On the Origin of Species:*

> It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method

> used in judging of the common events of life, and has often been used by the greatest natural philosophers. (Darwin 1872, 477)

In contrast to Peirce, several philosophers and scientists have argued that explanatory considerations, which make sense of confusing observations, can provide good reason to accept a hypothesis as true. Philosophers today discuss the possibility that the ability to explain a confusing set of observations can justify the acceptance of that explanation as true under the heading of "inference to the best explanation" (Harman 1965; Lipton 2004; Mohammadian 2021; Cabrera 2022). Although it has been common to equate IBE with abduction (Williamson 2018; Douven 2022), many have objected to this practice (Campos 2011; Mcauliffe 2015; Yu and Zenker 2018). Here, I take abduction to be concerned with the production and recommendation of hypotheses for investigation, IBE is concerned with their justification, or at least with the degree to which they are confirmed. Indeed, it is not clear that hypotheses should be entertained for the same reasons as they are selected as best (Henderson 2022). Peirce, for example, argued that fruitfulness or "uberty" should be prized over likelihood or security in abductions: "Peirce favored uberty over security in abductions, because likelihoods are often misleading guides that lead us to overlook potentially fertile hypotheses" (Mcauliffe 2015, 304). Those who doubt this should recall that once it was considered extremely unlikely that the Earth rotated about its axis, even though this hypothesis turned out to be most fruitful. The criteria for a good abduction might be different to that for the best explanation. Even so, the reasons for suggesting and further investigating a hypothesis may have much to do with the reasons for believing that the hypothesis is true. "Thus, it is not clear that abduction as a method of generation or a motive for pursuit can always be sharply separated from considerations of justification" (Niiniluoto 2018, 13). Nevertheless, here it is useful to distinguish suggestion from justification. Without endorsing the view that discovery and justification are strictly distinct, here abductions are explanation-based inferences that recommend plausible hypotheses for consideration, and IBEs are explanation-based inferences that justify the acceptance of hypotheses as true.

Like abduction IBE refers to different things. These include basic or fundamental inferences, which cannot be reduced to a collection of other forms of inference, and more complex inferential processes that combine deductions, inductions, abductions, and other forms of inference. The term "inference to the best explanation" was introduced by Gilbert Harman (1965), who described it as a fundamental inference in which "one infers, from the premise that a given hypothesis would provide a 'better' explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true" (1965, 89). More recently, others have characterized it as follows:

> The surprising fact, C, is observed;
> But if A were true, C would be a matter of course.
> No available competing hypothesis can explain C as well as A does.
> Hence, A is true. (Mackonis 2013, 977)

This formulation of IBE is similar to Peirce's formulation of abduction, except that the conclusion is that "A is true," rather than "there is reason to suspect that A is true," and that there is an extra premise about competing hypotheses. In this formulation, it is this extra

premise that warrants the stronger conclusion (Henderson 2022). Although Peirce held that abduction included the comparison of several competing hypotheses to determine which should be pursued further, IBE is usually characterized as a two-stage process in which a number of competing hypotheses are assembled in the first stage, before being compared in the second stage to determine which is the best explanation of the observed facts and thus be accepted as true (Dellsén 2021).

This characterization of IBE raises a few important questions. How should we judge which is the best explanation? Philosophers have suggested a number of "explanatory virtues," such as providing a mechanism, simplicity, explanatory scope (how many signs does the hypothesis explain?), explanatory power (does the hypothesis make the surprising unsurprising?), unifying power (does it bring together disparate domains of knowledge?), and fit with background knowledge (does it agree with what is already accepted as knowledge?) (Thagard 1978; Okasha 2000; Lipton 2004; Beebe 2009; Dragulinescu 2016; Schupbach 2017; Dellsén 2021). Some have argued that the best explanation should ideally be the only available hypothesis, or at least be much better than its competitors (Bird 2010). How to balance these virtues is a difficult issue that requires close consideration, as is the question of whether "explanatoriness" tells us anything about truth at all (Roche and Sober 2013; Cabrera 2017; McCain and Poston 2017; Schupbach 2017).

Apart from the question of which of the available explanations is best, how do we know that the true explanation is among the ones we consider? We might have dismissed the truth as implausible before comparing it to others, or worse, we might not even have thought of it at all. The explanation we select might be the best available, but still be false because it is "the best of a bad lot" (Van Fraassen 1989). In place of Peirce's faith in humanity's guessing instinct, others have suggested that our background knowledge, things that are already accepted as true, will help us to generate and deem plausible a set of hypotheses that include the truth (Lipton 2004; Schurz 2016; Dellsén 2021; Cabrera 2022). This makes our confidence that an explanation is true dependent on our background knowledge, which is itself fallible, but this is the case for any ampliative inference, not just IBE: "Confirmation is a three-place relationship between hypotheses, observations and background assumptions" (Sober 1988, 59).

Additionally, philosophers have argued that an explanation should not only be the best available, but also should be "satisfactory" or "good enough" (Musgrave 1988; Lipton 2004; Dellsén 2021). These, too, have been difficult notions to flesh out. Finnur Dellsén (2021) has argued that, as time goes on, new observations that support or fail to refute the best available explanation, as well as repeated failures to come up with a superior explanation, can eventually justify the view that this is not only the best available but also the best possible explanation. According to Dellsén (2021), this "explanatory consolidation" is not an instantaneous event, but rather a temporally extended process. This makes IBE a three-stage process involving the collection of candidate hypotheses, the selection of the best available hypothesis, and its explanatory consolidation as the best possible hypothesis. Furthermore, on this view, IBE would not be a basic or fundamental inference, but rather a complex inferential process composed of inductions, deductions, and explanation-based reasoning.

A prominent account of IBE is provided by Peter Lipton (2004), who also describes a complex inferential process modeled on his analysis of Ignaz Semmelweis's work. Plausible hypotheses are put forward in a first explanation-based stage. Observations that would

discriminate between these candidate hypotheses are deduced and then made, narrowing the field further by weeding out unsuccessful hypotheses. Another round of explanation-based reasoning is then used to rank the remaining hypotheses, and the one with the best explanation is chosen (Mohammadian 2021). Thus, IBE is described as both a basic method of inference and as a complex procedure involving many other forms of inference as well. Just how explanatory considerations inform the justification of hypotheses is a complex topic, which deserves close consideration by those who wish to make an IBE.

The explanations offered in this early shaken baby syndrome literature have many of the explanatory virtues that characterize IBE. Shaking explains a very confusing set of observations, drawing together observations from different fields of research (pediatric medicine, pathology, and experiments on primates). A mechanism by which shaking can cause the relevant pathology is provided. Shaking is said to provide a much better explanation than some alternatives, such as battering. However, it is difficult to frame the arguments made by early shaken baby syndrome researchers as compelling inferences to the best explanation.

For one thing, neither Kempe et al. (1962) nor Guthkelch (1971) suggested that abuse should be accepted as the best possible diagnosis in every patient with unexplained intracranial and retinal hemorrhage. They only argue that abuse should be entertained as a possible explanation. Kempe at al. (1962), for example, only concluded that abuse "should be considered" in such cases. Guthkelch (1971) also offered a rather tentative conclusion, to "keep in mind the possibility of assault" when presented with unexplained intracranial bleeding. More recently, Guthkelch argued that the sorts of observations these early researchers made should not have been used to infer that all such patients had been abused:

> While these events may have triggered a subdural hemorrhage or rehemorrhage, it is unwarranted to go from this possibility to the assumption that unexplained subdural hemorrhages, with or without retinal hemorrhage or encephalopathy, are caused by violent shaking or other forms of abuse. (2011, 206)

Caffey, however, did suggest that shaking could well be the explanation for all cases of unexplained intracranial and ocular bleeding. Caffey said that shaking "appears to be the major cause in these infants who suffer from subdural hematomas and intraocular bleedings" (1974, 402), and he suggested that a series of patients with unexplained intracranial and ocular bleeding may well have all been assaulted. Rather than being one possible explanation, abuse is offered by Caffey as the most common explanation for these pathological signs. Even Caffey, however, concluded that the state of the evidence as he had presented it was "manifestly incomplete and circumstantial" (1974, 403). Thus, Caffey might be read as making an IBE, but one that only lends a degree of confirmatory support to its conclusion, rather than putting it beyond a reasonable doubt.

Furthermore, while Caffey does compare the explanation of abusive shaking to other explanations, including battering and idiopathic retinal hemorrhage, he does not compare abusive shaking to all available hypotheses. Indeed, Caffey (1972, 165–166) suggested that the vibration from baby bouncers, driving on a rough road, or bouncing the child on the parent's knee, could damage an infant's brain: "There are several, apparently innocent, accepted, habitual practices, other than intentional shaking and jerking, which whiplash the head and brain, and which could lead to permanent brain damage" (Caffey 1972, 165). Thus,

Caffey did consider accidental trauma as an alternative explanation to abuse, but he provides no reason why abuse alone should be considered the superior explanation to the combination of abuse and accidental trauma. Caffey did compare some available and apparently plausible explanations, but not all of them.

Even if abusive shaking was the only explanation that researchers could think of, it would still be important to consider the possibility of unconceived alternatives. These have been a real problem in many areas of science (Stanford 2006), and a particular problem in medicine. In medicine, researchers always have to worry that a set of symptoms and signs has more than one explanation, as different diseases with different etiologies can present in the same way. For example, in the early modern period, "dropsy," a massive accumulation of body water in a patient who therefore presents as hugely edematous, was thought to be one disease. Today, doctors do not view such edematous patients as necessarily having the same disease. Doctors recognize that patients with heart failure, kidney failure, and liver diseases can all become edematous, but for different reasons. In medicine, the possibility that the same set of symptoms and signs might be caused by unrecognized conditions should always be taken seriously. What is more, the general rule that patients with unexplained intracranial and ocular bleeding have been abused is used to convict people of serious crimes and to remove children from parents. The criminal standard of evidence is beyond a reasonable doubt, and advocates of this rule claim that this standard has been met. For an explanation-based inference to be compelling in this context, it would need to provide arguments that the explanations considered are not just the *available* explanations, but also the *possible* explanations (Jellema 2021). The early shaken baby syndrome literature does not try to address the possibility unconceived alternatives.

Additionally, the arguments made by these early researchers were not the complex inferences that some philosophers use to describe IBE. Kempe et al. (1962) had suggested that patients with intracranial and ocular bleeding with particular fractures to their upper arm bones may be battered infants. Caffey (1972, 1974), synthesizing 25 years of work into these strange fractures, argued that shaking was a better explanation than battering in these infants. He argued that shaking was the best explanation for how a violent assault could produce intracranial injuries without injuring the outside of the head. One might argue that this provides some explanatory consolidation for what caused these fractures. Even if this is accepted, it does not show that shaking is the best explanation for all patients with unexplained intracranial and ocular bleeding. As Caffey was the first to make this claim, there was not time for a process of explanatory consolidation to take place in the early 1970s. Caffey's work (1972, 1974) does not describe the computation and collection of observations that could distinguish between competing hypotheses. The early research into shaken baby syndrome does not fit the description of IBE as a complex inference offered in some contemporary philosophical accounts.

In summary, the early shaken baby syndrome literature offers explanation-based inferences, which support their conclusions somewhat tentatively, without comparing all available hypothesis to establish which one is best. The problem of unconceived alternatives is not discussed, and the arguments deployed are not the complex inferences that some say characterize IBE. Consequently, I suggest it is better to regard these arguments as abductions, rather than as inferences to the best explanation.

Even so, some of the language used by these early shaken baby researchers is ambiguous, and can be read as claiming that unexplained intracranial and ocular bleeding

is almost always the result of violent assault. Aspects of this presentation are described as "a major diagnostic feature," the "most characteristic pattern of physical findings," and "the essential elements" of battered, whiplashed or shaken infants (Kempe et al. 1962, 18; Caffey 1974, 399, 402). These claims can be read in different ways. On the one hand, they might only mean that these observations have been made in many cases of assaulted infants. However, as to be diagnostic, to be characteristic, and to be part of something's essence can be understood as being unique to that thing, these claims might also be read to mean that these signs had been observed only to occur in abused patients. This ambiguity in language makes it difficult to tell what has been observed and what is being claimed, and even blur the distinction between observing diseased patients to see what signs they have, and observing patients with signs to see what diseases they have.

Perhaps this impression is amplified by naming a syndrome, a collection of symptoms and signs, after a suspected explanation of that syndrome. This practice invites the confusion between understanding a syndrome as (1) a collection of symptoms and signs; and as (2) a collection of symptoms and signs explained in a particular way. Recognizing that there are patients who have been shaken and develop unexplained intracranial and ocular bleeding, and thus who have shaken baby syndrome (sense 2), might be interpreted as meaning that patients with unexplained intracranial and ocular bleeding have this same syndrome (sense 1), and thus must have been shaken. Does the claim "shaken baby syndrome exists" mean that there are patients who have been shaken who develop these pathological signs, or does it mean that patients with these pathological signs have all been shaken? This is why Guthkelch (2011, 202) recommended calling the syndrome "retino-dural haemorrhage of infancy," so as to remain neutral about what the explanation of this presentation is. How we use language can give the impression that a certain presentation has been observed to be unique to a particular condition, without anyone actually claiming that this is the case.

8. The Early Research Was Put to Work

Soon after the publication of this early research into shaken baby syndrome, other researchers used it to study shaken baby syndrome. For example, Ludwig and Warman (1984) studied twenty cases that they identified as shaken infants. These researchers coined the phrase "shaken baby syndrome" (Christian and Block 2009) and provide an early example of work that uses the conclusions of the early shaken baby research to make their own observations. Ludwig and Warman (1984) identified cases with subdural and retinal hemorrhage whose parents had admitted shaking them, or whose parents provided a history that (in their view) could not explain this intracranial and intraocular bleeding. They then excluded cases that had evidence of external injuries to the head, because they only wanted to study infants who were shaken, and not infants who were battered.

> Cases of possible shake injury were selected based on parent's admission of shaking or suspicion by medical staff of this form of abuse when the history and evaluation could not account for the patient's injuries.
>
> Only children injured solely by being shaken were included in the study. Therefore, patients with other evidence of abuse (external head trauma, skull fracture, multiple

> skeletal fractures, burns, or patterned or severe bruising) were excluded despite the possibility of their having concomitant shake injuries (Ludwig and Warman 1984, 104).

Only in three of these cases did the parents admit to shaking, and only then as part of an effort to resuscitate an already very sick infants. "Rather than a history of abuse, most parents presented the problem in terms of respiratory abnormality" (Ludwig and Warman 1984, 105). In eight cases the parents described a recent accidental injury, which Ludwig and Warman (1984) apparently did not believe could have caused the intracranial and intraocular bleeding. Ludwig and Warman (1984) selected their cases as if all patients with unexplained intracranial and ocular bleeding *must* have been shaken. Using this general rule, they deduced that patients with this presentation had been violently assaulted. They argued that all infants with unexplained intracranial and ocular bleeding have been assaulted, these infants have unexplained intracranial and ocular bleeding; therefore, these patients have been assaulted.

To make such a deduction requires a very high level of confidence that this general rule is true. As discussed, the early shaken baby syndrome literature provided an abduction, a fair guess, that this might be the case. This is hardly sufficient to warrant the use of this general rule to deduce that an infant has been assaulted. To justify this rule, Ludwig and Warman (1984, 104) cited Caffey's papers (1972, 1974), also saying that "other authors have confirmed Caffey's observations and the likely mechanism and pathogenesis of the injuries." Thus, this further research might have taken Caffey's suggestion and confirmed it by making a really compelling IBE. If sufficiently compelling, this might justify the high level of confidence in this general rule held by Ludwig and Warman (1984).

However, it is difficult to see how this further research, even in conjunction with Caffey's work, could be interpreted as supporting a compelling IBE. The further research provides a series of case studies in which shaken infants are observed to develop unexplained intracranial and ocular bleeding (Zimmerman et al. 1979; Bennett and French 1980; McClelland et al. 1980).[10] The possibility of alternative explanations for these pathological signs is not discussed in any depth, let alone dismissed as impossible. One study reports that skull fractures were more common in non-abused infants, and that particular kind of subdural hemorrhage was rarely found in non-abused children with head trauma, but does not discuss non-abused infants further (Zimmermann et al. 1979). Providing more and more instances of patients who have been abused who develop these pathological signs does very little, if anything, to support the claim that all patients with these pathological signs must have been abused, any more than observing more and more ripe bananas that are yellow things supports the view that all yellow things must be ripe bananas. And yet, Ludwig and Warman (1984) appear to have thought that it does. In effect, they tried to use observations of patients *with* a disease to show that certain pathological signs are specific for that disease, when specificity can only be assessed by observing patients *without* the disease. The failure to find patients with unexplained intracranial and ocular bleeding in patients who definitely had not been abused despite a thorough search would have been much more compelling.

[10] Note that many of these additional cases were identified by assuming that patients with unexplained intracranial and ocular bleeding have been abused, as Ludwig and Warman (1984) did. This assumption was used to identify cases that made Ludwig and Warman more confident that this assumption was true, so they used it to identify more cases. In this way, we can see how circular arguments might have entered this literature.

Perhaps it is possible that Ludwig and Warman (1984) confused abduction with IBE. The observation that there are patients with a disease who have certain pathological signs provides good grounds for the abduction that patients with those signs might have that disease. However, a good abduction need not be a good IBE. If a good abduction is mistaken for a good IBE, the conclusion may seem more secure than it should be. The ambiguous language identified may make this more likely by producing confusion about what the conclusion is. An abduction from those cases would be: there are patients with surprising pathological signs; if patients had this disease, these signs would occur as a matter of course; therefore, there is reason to suspect that the patients have this disease. If the conclusion of this argument is taken to be much stronger, as it might be in a really compelling IBE, the argument would be: there are patients with surprising pathological signs; if patients had this disease, these signs would occur as a matter of course; therefore, the patients have this disease. The form of this argument is similar to the logical fallacy of affirming the consequent.[11] Ripe bananas are yellow things; therefore, yellow things are ripe bananas. If this is what happened in this case, abduction and IBE need to be more carefully distinguished.

Alternatively, Ludwig and Warman (1984) might have assumed that there can be only one explanation for unexplained intracranial and ocular bleeding. I argued above that this would be a foolish assumption, particularly in medicine. Even so, if they assumed this, finding *an* explanation for these surprising phenomena would be the same as finding *the* explanation for them. However, if this was the case, establishing that this happened in a single case would be adequate to establish that this was the correct explanation. Finding more and more cases that could be explained like this would be superfluous.

Finding more and more cases would not be superfluous, however, if the general rule that all patients with these pathological signs have been violently assaulted was arrived at by induction. If Caffey and others advanced the view that patients with unexplained intracranial and ocular bleeding have always been assaulted, and then other researchers investigated patients who have these pathological signs and found that they had indeed been assaulted, this might be compelling evidence that this claim was true. The more times this is found to be the case, the more compelling this argument becomes. However, this argument is a more-of-the-same inference: it is an induction, not an abduction. In order to establish that patients with certain signs always have a disease by induction, observations of patients with those signs who have that disease are required.

It would be easy to suppose that observations of this kind were indeed provided by the researchers cited by Ludwig and Warman (1984) and by Caffey (1972, 1974). All the cases presented by these researchers were patients with unexplained intracranial and ocular bleeding who had been abused. Thus, it seems fair to say that researchers did observe patients with these pathological signs who had been assaulted and that they made an induction. However, appearances can be deceiving. We must not forget Figure 1. These cases are not a fair sample of cases from the population of patients with these signs—the green circle in Figure 1. They were chosen because they were cases with both the pathological signs and the disease—they were from the red area in Figure 1. Any attempt to present these cases as a fair sample of patients with those pathological signs would introduce an egregious

[11] Affirming the consequent: if *a,* then *b; b,* therefore, *a*. There are several other analogous accounts of affirming the consequent, regarding fetal alcohol spectrum disorder, Alzheimer's disease and ADHD (Tait 2009; Price and Miskelly 2015; Herrup 2022).

sampling error and be an instance of confirmation bias. If framed as an induction, it would be a very poor induction indeed. Even so, believing that the general rule is supported by induction would be an easy mistake to make. Ludwig and Warman might have confused the observations that are useful for making an abduction with observations that are useful for making an induction. Indeed, ambiguous language may have made it difficult to distinguish these observations, or even to understand that they are different. Perhaps this is why Ludwig and Warman (1984) think finding more and more these cases of abuse and unexplained intracranial and ocular bleeding makes Caffey's conclusions highly compelling.

If Ludwig and Warman (1984) were not so utterly confident that patients with this presentation had been shaken, their results might have given them pause. They argue that meningitis is a poor explanation for retinal hemorrhage in their cases, but that is all. Why is it that so many of their cases presented with histories of breathing difficulties? Might it be because these are part of another hitherto unanticipated explanation for this bleeding? Or might it be because breathing difficulties and this bleeding all share a hitherto unanticipated causal factor? Why did so many of their cases, in addition to the ones reported by Caffey (1974), have no history of trauma? Could it be that some of these parents were telling the truth? Caffey (1974) suggested that minor trauma from rough handling could produce intracranial bleeding. Could it be that the reported minor accidents caused the bleeding? Ludwig and Warman (1984) do not even discuss these possibilities. They don't appear interested in how these difficult-to-explain observations might challenge their belief that patients with unexplained intracranial and ocular bleeding have been violently assaulted. If they had understood the observations they refer to as supporting an abduction, a fair guess, this would have been incautious. However, if they understood these observations as supporting a highly compelling induction or IBE, their faith in this belief is easier to understand.

This same suspicion is raised again by the interpretation of biomechanical evidence in a paper by Duhaime and colleagues (1987). This study is important for introducing the concept that shaking alone might not be sufficient to produce intracranial and retinal bleeding, and that impact might be necessary as well (Reece 2001; Cory and Jones 2003; Christian and Block 2009). These researchers built a mechanical model of an infant child, similar to a crash test dummy, to measure the acceleration of the infant brain that could be produced by violent shaking. They found that even the most forceful shaking that strong people could inflict did not approach the estimated thresholds required to produce intracranial bleeding. They also found, surprisingly, that impact to the head from short falls *did* breach this threshold. Accepting the biofidelity of their model, they concluded that shaking alone could not cause intracranial bleeding: "It is our conclusion that the shaken baby syndrome, at least in its most severe and acute form, is not usually caused by shaking alone" (Duhaime et al. 1987, 414). However, they did not question the view that children with these signs had been assaulted:

> Although shaking may, in fact, be part of the process, it is more likely that such infants suffer blunt impact. The most common scenario may be a child who is shaken, then thrown into or against a crib or other surface, striking the back of the head and thus undergoing a large, brief deceleration. (Duhaime et al. 1987, 414)

Duhaime and colleagues (1987) also presented a case series of violently assaulted patients, which they identified by accepting that all infants with unexplained intracranial and ocular bleeding have been violently assaulted, as Ludwig and Warman (1984) had done (Duhaime et al. 1987, 410). If it was reached by IBE, the conclusion that (just about) all patients with unexplained intracranial and ocular bleeding have been violently assaulted is dependent upon shaking being an excellent explanation for these signs. Undermining shaking as an explanation has the power to undermine this conclusion as well.

Inductions are not always vulnerable to changes to the perceived quality of a proposed explanation. For example, doctors have been using general anesthesia for over a century without being able to explain how general anesthetics produce unconsciousness (Mashour, Forman, and Campagna 2005; Gent and Adamantidis 2017). Confidence in the ability to induce and control anesthesia is based on the correlation between the use and dose of a drug and its effect on consciousness. If new evidence was produced that undermined one of the candidate explanations for how general anesthetics work, this would not undermine confidence in the drug's ability to induce anesthesia. This is because this confidence is based on the correlation of the drug's use and its effect, and not on the quality of the explanation of how it works.

Nevertheless, Duhaime and colleagues' (1987) confidence in their conclusion remained undiminished. They responded to the new evidence as if the general rule that all children with unexplained intracranial and ocular bleeding had been violently assaulted was an observation, supported by a compelling induction, which needed to be explained somehow, even if shaking alone could not account for it. If they understood this general rule as supported by an abduction, undermining the explanation of how abuse produces these pathological signs might have undermined their faith in this relationship itself.

9. Conclusion

Paying attention to the methods of inference used by researchers during the development of the medical literature on shaken baby syndrome is useful. As mentioned earlier, the contemporary literature on this issue has become acrimonious, with each side accusing the other of foolishness and bias. One way to move forward is to return to analysing evidence and argument, especially by thinking about the methods of inference used to support conclusions.

A central claim made by advocates of the traditional view of shaken baby syndrome is that unexplained intracranial and ocular bleeding are tightly correlated with abuse, such that this presentation is almost always the result of a violent assault. Opponents of the traditional view hold that no such correlation has ever been established. Instead, they say, the medical literature on shaken baby syndrome is rife with circular arguments. As medical and legal decisions about whether an infant has been abused are made using the belief that patients with unexplained intracranial and ocular bleeding have very likely been assaulted, it is inevitable that this presentation will only occur in patients considered to have been abused—regardless of whether they have actually been abused. This belief does not have an empirical origin at all, but is rather a sort of dogma with which pediatricians have become indoctrinated.

Quite reasonably, advocates of the traditional view object to this characterization. The belief that unexplained intracranial and ocular bleeding is almost always the result of a

violent assault cannot entirely be the result of a circular argument. At some point in time researchers believed no such thing, and had to discover that this was the case by observing their patients. This belief must have an empirical basis somewhere.

By paying attention to the methods of inference used in the early shaken baby syndrome literature, I suggest that we can see what happened here. This belief does indeed have an empirical basis, but it is not the result of observing a correlation. This belief was not the result of an *induction,* but rather of an *abduction*. Early shaken baby researchers observed that patients who had been violently assaulted were patients with unexplained intracranial ocular bleeding and they used this observation to infer that patients with unexplained intracranial and ocular bleeding were patients who had been violently assaulted. Rather than inferring that patients with those signs have this disease from the observation that patients with those signs have this disease (which would be an induction), these early researchers inferred that patients with those signs have this disease from observations that patients with that disease have those signs (by abduction). These early researchers argued that this explanation could make sense of the confusing observation of unexplained intracranial and ocular bleeding, and thus to suspect that this explanation is correct.

I have argued that the arguments made by early researchers should be thought of as abductions—as a fair guess about what might be going on with these patients. They should not be thought of as compelling inferences to the best explanation, which are used to try to justify a belief. Nevertheless, some researchers quickly accepted the suggestions made by the early shaken baby researchers as unassailable conclusions. This might be because they mistook a perfectly good abduction for a compelling IBE. If so, this would resemble the fallacy of affirming the consequent. Alternatively, they may have confused a perfectly good abduction with a compelling induction. This confusion might lead them to be more confident in their conclusions than they otherwise would be. The method of inference a researcher thinks is being used to draw a conclusion can influence the degree of confidence they have in their conclusion and should be considered closely. A perfectly good abduction should not be confused with a compelling IBE or a compelling induction.

Ludwig and Warman (1984) were so confident of their ability to diagnose abuse in patients with unexplained intracranial and ocular bleeding that they deduced that such patients must have been abused. Patients with these pathological signs would be recorded as abused, even if they were not, making it appear that these pathological signs and abuse are tightly correlated, even if they are not. If others did the same thing when diagnosing abuse, it is easy to see how circular arguments entered this literature. This belief does have an empirical origin, but it was the product of an abduction—a fair guess. Excessive confidence in the truth of this fair guess can produce circular arguments, which might have led to the illusion that this belief is firmly established by statistical correlation. I suggest that this may have happened in the shaken baby syndrome literature, but recognize that more investigation of this literature is necessary to see if this is the case

Thinking about methods of inference can also alter how researchers react to new evidence. Being based on explanatory power, abductions are vulnerable to evidence that undermines explanations in a way that inductions often are not. If researchers believe that the association between unexplained intracranial and ocular bleeding and abuse has been established by induction, they may believe that there is an empirical relationship that must be explained, even if an explanation is not forthcoming. However, if this association is suggested by abduction, the lack of a good explanation may be interpreted to undermine the

suggested association itself. The method of inference used to support claims matters and should be considered closely.

Readers may quite reasonably disagree with my interpretation of these arguments. It may be that the early shaken baby researchers made a highly compelling IBE, justifying the belief that patients with unexplained intracranial and ocular bleeding have just about always been abused. Even if the early literature only provided a fair guess at the truth, it may be that since then new evidence has emerged that allows such a highly compelling IBE to be made. If this is the case, the way forward in this debate is to analyse the evidence and arguments used to reach the conclusions held by people today. Rather than focusing on inductive arguments, it may be possible to make highly compelling inferences to the best explanation on this matter, but this will need to be done carefully and explicitly.

Acknowledgments

The research for this article was supported by the Dutch Research Council (Nederlandse Organisatie voor Wetenschappelijk Onderzoek—NWO) as part of the project "Health and Disease as Practical Concepts: A Pragmatist Approach to Conceptualization of Health and Disease" (project no. 406.18.FT.002). https://www.erasmusmc-rdo.nl/project/health-and-disease-as-practical-concepts/.

Disclosure Statement

No competing interest was reported by the author.

References

American Academy of Pediatrics: Committee on Child Abuse and Neglect. 2001. "Shaken Baby Syndrome: Rotational Cranial Injuries—Technical Report." *Pediatrics* 108, no. 1: 206–210. https://doi.org/10.1542/peds.108.1.206.

Beebe, James R. 2009. "The Abductivist Reply to Skepticism." *Philosophy and Phenomenological Research* 79, no. 3: 605–636. https://doi.org/10.1111/j.1933-1592.2009.00295.x.

Bennett, Harvey S. and Joseph H. French. 1980> "Elevated Intracranial Pressure in Whiplash-Shaken Infant Syndrome Detected with Normal Computerized Tomography." *Clinical Pediatrics* 19, no. 9: 633–634. https://doi.org/10.1177/000992288001900912.

Bird, Alexander. 2010. "Eliminative Abduction: Examples from Medicine." *Studies in History and Philosophy of Science Part A* 41, no. 4: 345–352. https://doi.org/10.1016/j.shpsa.2010.10.009.

Bolton, Jonathan W. 2015. "Varieties of Clinical Reasoning." *Journal of Evaluation in Clinical Practice* 21, no. 3: 486–489. https://doi.org/10.1111/jep.12309.

Cabrera, Frank. 2017. "Can There Be a Bayesian Explanationism? On the Prospects of a Productive Partnership." *Synthese* 194, no. 4: 1245–1272. https://doi.org/10.1007/s11229-015-0990-z.

———. 2022. "Inference to the Best Explanation: An Overview." In *Handbook of Abductive Cognition,* edited by Lorenzo Magnani, 1–34. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-68436-5_77-1.

Caffey, John. 1972. "On the Theory and Practice of Shaking Infants: Its Potential Residual Effects of Permanent Brain Damage and Mental Retardation." *American Journal of Diseases of Children* 124, no. 2: 161–169. https://doi.org/10.1001/archpedi.1972.02110140011001.

———. 1974. "The Whiplash Shaken Infant Syndrome: Manual Shaking by the Extremities with Whiplash-Induced Intracranial and Intraocular Bleedings, Linked with Residual Permanent Brain Damage and Mental Retardation." *Pediatrics* 54, no. 4: 396–403. https://doi.org/10.1542/peds.54.4.396.

Campos, Daniel G. 2011. "On the Distinction between Peirce's Abduction and Lipton's Inference to the Best Explanation." *Synthese* 180, no. 3: 419–442. https://doi.org/10.1007/s11229-009-9709-3.

Chiffi, Daniele. 2020. *Clinical Reasoning: Knowledge, Uncertainty, and Values in Health Care*. Cham: Springer.

Choudhary, Arabinda Kumar, Sandeep K. Narang, Joëlle Anne Moreno, Cindy W. Christian, Sabah Servaes, Vincent J. Palusci, Gary L. Hedlund, et al. 2019. "A Consensus Response on the Complete Picture: Reply to Lynøe and Eriksson." *Pediatric Radiology* 49, no. 3: 424–428. https://doi.org/10.1007/s00247-019-04341-4.

Choudhary, Arabinda Kumar, Sabah Servaes, Thomas L. Slovis, Vincent J. Palusci, Gary L. Hedlund, Sandeep K. Narang, Joëlle Anne Moreno, et al. 2018. "Consensus Statement on Abusive Head Trauma in Infants and Young Children." *Pediatric Radiology* 48, no. 8: 1048–1065. https://doi.org/10.1007/s00247-018-4149-1.

Christian, Cindy W. and Robert Block. 2009. "Abusive Head Trauma in Infants and Children." *Pediatrics* 123, no. 5: 1409–1411. https://doi.org/10.1542/peds.2009-0408.

Climenhaga, Nevin. Forthcoming. "Evidence and Inductive Inference." In *The Routledge Handbook of the Philosophy of Evidence,* edited by Maria Lasonen-Aarnio and Clayton Littlejohn. New York: Routledge. A preview of this chapter is available here: http://philsci-archive.pitt.edu/17450/1/Evidence%20and%20Inductive%20Inference.pdf.

Cohen, Marta C. 2019. "The Swedish Shaken Baby Syndrome Report and Review: What Does the Latest Knowledge Tell Us?" *Acta Paediatrica* 108, no. 2: 198–199. https://doi.org/10.1111/apa.14643.

Cory, C.Z. and M.D. Jones. 2003. "Can Shaking Alone Cause Fatal Brain Injury? A Biomechanical Assessment of the Duhaime Shaken Baby Syndrome Model." *Medicine, Science and the Law* 43, no. 4: 317–333. https://doi.org/10.1258/rsmmsl.43.4.317.

Darwin, Charles. 1872. *The Origin of Species*. London: John Murray.

Debelle, Geoffrey David, Sabine Maguire, Patrick Watts, Rosa Nieto Hernandez, and Alison Mary Kemp. 2018. "Abusive Head Trauma and the Triad: A Critique on Behalf of RCPCH of 'Traumatic Shaking: The Role of the Triad in Medical Investigations of Suspected Traumatic Shaking'." Archives of Disease in Childhood 103, no. 6: 606–610. https://doi.org/10.1136/archdischild-2017-313855.

Dellsén, Finnur. 2021. "Explanatory Consolidation: From 'Best' to 'Good Enough'." *Philosophy and Phenomenological Research* 103, no. 1: 157–177. https://doi.org/10.1111/phpr.12706.

Donohoe, Mark. 2003. "Evidence-Based Medicine and Shaken Baby Syndrome: Part I: Literature Review, 1966–1998." *American Journal of Forensic Medicine and Pathology* 24, no. 3: 239–242. https://doi.org/10.1097/01.paf.0000083635.85457.97.

Douven, Igor. 2022. *The Art of Abduction*. Cambridge, MA: Massachusetts Institute of Technology Press.

Dragulinescu, Stefan. 2016. "Inference to the Best Explanation and Mechanisms in Medicine." *Theoretical Medicine and Bioethics* 37, no. 3: 211–232. https://doi.org/10.1007/s11017-016-9365-9.

Duhaime, Ann-Christine and Cindy W. Christian. 2019. "Abusive Head Trauma: Evidence, Obfuscation, and Informed Management: JNSPG 75th Anniversary Invited Review Article." *Journal of Neurosurgery: Pediatrics* 24, no. 5: 481–488. https://doi.org/10.3171/2019.7.PEDS18394.

Duhaime, Ann-Christine, Thomas A. Gennarelli, Lawrence E. Thibault, Derek A. Bruce, Susan S. Margulies, and Randall Wiser. 1987. "The Shaken Baby Syndrome: A Clinical, Pathological, and Biomechanical Study." *Journal of Neurosurgery* 66, no. 3: 409–415. https://doi.org/10.3171/jns.1987.66.3.0409.

Fann, K.T. 1970. *Peirce's Theory of Abduction*. The Hague: Martinus Nijhoff.

Findley, Keith A., Patrick D. Barnes, David A. Moran, and Wayne Squier. 2012. "Shaken Baby Syndrome, Abusive Head Trauma, and Actual Innocence: Getting it Right." *12 Houston Journal of Health Law and Policy* 209, Univ. of Wisconsin Legal Studies Research Paper No. 1195. https://ssrn.com/abstract=2048374.

Findley, Keith A. Michael Risinger, Patrick Barnes, Julie Mack, David A. Moran, Barry Scheck, and Thomas Bohan. 2019. "Feigned Consensus: Usurping the Law in Shaken Baby Syndrome/Abusive Head Trauma Prosecutions." *Wisconsin Law Review* 2019, no. 5: 1211–1268. https://dx.doi.org/10.2139/ssrn.3328996.

Flach, Peter A. and Antonis C. Kakas. 2000. "Abductive and Inductive Reasoning: Background and Issues." In *Abduction and Induction: Essays on their Relation and Integration,* edited by Peter A. Flach and Antonis C. Kakas, 1–27. Dordrecht: Springer Netherlands (Applied Logic Series). https://doi.org/10.1007/978-94-017-0606-3_1.

Gabaeff, Steven C. 2016. "Exploring the Controversy in Child Abuse Pediatrics and False Accusations of Abuse." *Legal Medicine* 18: 90–97. https://doi.org/10.1016/j.legalmed.2015.12.004.

Gent, Thomas and Antoine Adamantidis. 2017. 'Anaesthesia and Sleep: Where Are We Now?" *Clinical and Translational Neuroscience* 1, no. 1. https://doi.org/10.1177/2514183X17726281.

Guthkelch, A.N. 1971. "Infantile Subdural Haematoma and Its Relationship to Whiplash Injuries." *British Medical Journal* 2, no. 5759: 430–431. https://doi.org/10.1136/bmj.2.5759.430.

———. 2011. "Problems of Infant Retino-Dural Hemorrhage with Minimal External Injury." *Houston Journal of Health Law & Policy* 12, no. 2: 201–208. https://law.uh.edu/hjhlp/volumes/Vol_12_2/Guthkelch.pdf.

Hanson, Norwood Russell. 1958. *Patterns of Discovery: An Inquiry Into the Conceptual Foundations of Science*. Cambridge: Cambridge University Press.

Harman, Gilbert H. 1965. "The Inference to the Best Explanation." *Philosophical Review* 74, no. 1: 88–95. https://doi.org/10.2307/2183532.

Hellgren, Kerstin, Ann Hellström, Anna-Lena Hård, Lena Jacobson, Ulrika Lidén, Stefan Löfgren, Kristina Teär Fahnehjelm, and Jan Ygge. 2017. "The New Swedish Report on Shaken Baby Syndrome Is Misleading." *Acta Paediatrica* 106, no. 7: 1040. https://doi.org/10.1111/apa.13845.

Henderson, Leah. 2020. "The Problem of Induction." In *The Stanford Encyclopedia of Philosophy,* edited by Edward N. Zalta. Spring 2020. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2020/entries/induction-problem/.

———. 2022. "Putting Inference to the Best Explanation into Context." *Studies in History and Philosophy of Science* 94: 167–176. https://doi.org/10.1016/j.shpsa.2022.06.009.

Herrup, Karl. 2022. "Fallacies in Neuroscience: The Alzheimer's Edition." *eNeuro* 9, no. 1. https://doi.org/10.1523/ENEURO.0530-21.2021.

Högberg, Ulf, Erik Lampa, Göran Högberg, Peter Aspelin, Fredrik Serenius, and Ingemar Thiblin. 2018. "Infant Abuse Diagnosis Associated with Abusive Head Trauma Criteria: Incidence Increase Due to Overdiagnosis?" *European Journal of Public Health* 28, no. 4: 641–646. https://doi.org/10.1093/eurpub/cky062.

Ingraham, Franc D. and Donald D. Matson. 1954. *Neurosurgery of Infancy and Childhood.* Springfield: Charles C. Thomas.

Janson, Staffan. 2020. "Towards a Deeper Understanding of Abusive Head Trauma." *Acta Paediatrica* 109, no. 7: 1290–1291. https://doi.org/10.1111/apa.15164.

Jellema, Hylke. 2021. 'The Reasonable Doubt Standard as Inference to the Best Explanation." *Synthese* 199, no. 1: 949–973. https://doi.org/10.1007/s11229-020-02743-8.

Johannesson, Eric. 2022. "The Statistical Riddle of Induction." *Australasian Journal of Philosophy:* 1–14. https://doi.org/10.1080/00048402.2021.2013909.

Karlsen, Bjørg, Torgeir Martin Hillestad, and Elin Dysvik. 2021. "Abductive Reasoning in Nursing: Challenges and Possibilities." *Nursing Inquiry* 28, no. 1. https://doi.org/10.1111/nin.12374.

Kempe, C. Henry, Frederic N. Silverman, Brandt F. Steele, William Droegemueller, and Henry K. Silver. 1962. "The Battered-Child Syndrome." *JAMA* 181, no. 1: 17–24. https://doi.org/10.1001/jama.1962.03050270019004.

Laurent-Vannier, Anne, Catherine Adamsbaum, Jean-Sébastien Raul, Caroline Rey-Salmon, and Caroline Rambaud. 2018. "Flawed Swedish Study on Traumatic Shaking Is Already Being Used by Defence Lawyers and Its Findings Must Be Ignored." *Acta Paediatrica* 107, no. 12: 2048–2050. https://doi.org/10.1111/apa.14564.

Lindberg, Daniel M., Howard Dubowitz, Randell C. Alexander, and Robert M. Reece. 2019. "The 'New Science' of Abusive Head Trauma." *International Journal on Child Maltreatment: Research, Policy and Practice* 2, no. 1–2: 1–16. https://doi.org/10.1007%2Fs42448-019-00021-w.

Lipton, Peter. 2004. *Inference to the Best Explanation.* 2nd edition. London: Routledge. https://doi.org/10.4324/9780203470855.

Lucas, Steven, Anna Bärtås, Anna-Karin Edstedt Bonamy, Lisa Törnudd, Peter Wide, and Gabriel Otterman. 2017. "The Way Forward in Addressing Abusive Head Trauma in Infants: Current Perspectives from Sweden." *Acta Paediatrica* 106, no. 7: 1033–1035. https://doi.org/10.1111/apa.13840.

Ludwig, Stephen and Matt Warman. 1984. "Shaken Baby Syndrome: A Review of 20 Cases." *Annals of Emergency Medicine* 13, no. 2: 104–107. https://doi.org/10.1016/S0196-0644(84)80571-5.

Lynøe, Niels, Göran Elinder, Boubou Hallberg, Måns Rosén, Pia Sundgren, and Anders Eriksson. 2017a. 'Authors' Overarching Reply to All the Responses Received to the Systematic Literature

Review on Shaken Baby Syndrome." *Acta Paediatrica* 106, no. 7: 1031. https://doi.org/10.1111/apa.13887.

———. 2017b. "Insufficient Evidence for 'Shaken Baby Syndrome': A Systematic Review." *Acta Paediatrica* 106, no. 7: 1021–1027. https://doi.org/10.1111/apa.13760.

Lynøe, Niels and Anders Eriksson. 2019. 'The Swedish Systematic Review of Traumatic Shaking Is Not Flawed and Should Not Be Ignored." *Acta Paediatrica* 108, no. 2: 381. https://doi.org/10.1111/apa.14575.

Lynøe, Niels, Måns Rosén, Göran Elinder, Boubou Hallberg, Pia Sundgren, and Anders Eriksson. 2018. "Pouring Out the Dirty Bathwater without Throwing Away Either the Baby or Its Parents: Commentary to Saunders et al." *Pediatric Radiology* 48, no. 2: 284–286. https://doi.org/10.1007/s00247-017-4003-x.

Mackonis, Adolfas. 2013. "Inference to the Best Explanation, Coherence and Other Explanatory Virtues." *Synthese* 190, no. 6: 975–995. https://doi.org/10.1007/s11229-011-0054-y.

Magnani, Lorenzo. 2004. "Model-Based and Manipulative Abduction in Science." *Foundations of Science* 9, no. 3: 219–247. https://doi.org/10.1023/B:FODA.0000042841.18507.22.

———. 2009. *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*. Cham: Springer Science & Business Media.

Maguire, Sabine Ann, Alison Mary Kemp, Rebecca Caroline Lumb, and Daniel Mark Farewell. 2011. "Estimating the Probability of Abusive Head Trauma: A Pooled Analysis." *Pediatrics* 128, no. 3: e550–e564. https://doi.org/10.1542/peds.2010-2949.

Mashour, George A., Stuart A. Forman, and Jason A. Campagna. 2005. "Mechanisms of General Anesthesia: From Molecules to Mind." *Best Practice & Research Clinical Anaesthesiology* 19, no. 3: 349–364. https://doi.org/10.1016/j.bpa.2005.01.004.

Matschke, Jakob, Bernd Herrmann, Jan Sperhake, Friederike Körber, Thomas Bajanowski, and Markus Glatzel. 2009. "Shaken Baby Syndrome." *Deutsches Ärzteblatt International* 106, no. 13: 211–217. https://doi.org/10.3238/arztebl.2009.0211.

Mcauliffe, William H.B. 2015, "How Did Abduction Get Confused with Inference to the Best Explanation?" *Transactions of the Charles S. Peirce Society* 51, no. 3: 300–319. https://doi.org/10.2979/trancharpeirsoc.51.3.300.

McCain, Kevin and Ted Poston. 2017. *Best Explanations: New Essays on Inference to the Best Explanation*. Oxford: Oxford University Press.

McClelland C.Q., H. Rekate, B. Kaufman, and L. Persse. 1980. "Cerebral Injury in Child Abuse: A Changing Profile." *Pediatric Neurosurgery* 7, no. 5: 225–235. https://doi.org/10.1159/000119950.

Mohammadian, Mousa. 2021. "Abduction: The Context of Discovery + Underdetermination = Inference to the Best Explanation." *Synthese* 198, no. 5: 4205–4228. https://doi.org/10.1007/s11229-019-02337-z.

Musgrave, Alan. 1988. "The Ultimate Argument for Scientific Realism." In *Relativism and Realism in Science,* edited by Robert Nola, 229–252. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-2877-0_10.

Narang, Sandeep. 2011. "A *Daubert* Analysis of Abusive Head Trauma/Shaken Baby Syndrome." *Houston Journal of Health Law and Policy* 11: 505–633. https://www.law.uh.edu/hjhlp/volumes/Vol_11_3/Narang.pdf.

Narang, Sandeep K., Cynthia Estrada, Sarah Greenberg, and Daniel Lindberg. 2016. "Acceptance of Shaken Baby Syndrome and Abusive Head Trauma as Medical Diagnoses." *Journal of Pediatrics* 177: 273–278. https://doi.org/10.1016/j.jpeds.2016.06.036.

Niiniluoto, Ilkka. 2018. *Truth-Seeking by Abduction*. Cham: Springer. https://link.springer.com/book/10.1007/978-3-319-99157-3.

Okasha, Samir. 2000. "Van Fraassen's Critique of Inference to the Best Explanation." *Studies in History and Philosophy of Science Part A* 31, no. 4: 691–710. https://doi.org/10.1016/S0039-3681(00)00016-9.

Ommaya, Ayub K., Fred Faas, and Philip Yarnell. 1968. "Whiplash Injury and Brain Damage: An Experimental Study." *JAMA* 204, no. 4: 285–289. https://doi.org/10.1001/jama.1968.03140170001001.

Ommaya, Ayub K. and Philip Yarnell. 1969. "Subdural Haematoma after Whiplash Injury." *Lancet* 2, no. 7614: 237–239. https://doi.org/10.1016/s0140-6736(69)90005-1.

O'Rorke, P., S. Morris, and D. Schulenburg. 1990. "Theory Formation by Abduction: A Case Study Based on the Chemical Revolution." In *Computational Models of Scientific Discovery and Theory Formation,* edited by Jeff Shrager and Pat Langley, 197–224. San Mateo, CA: Morgan Kaufman.

Paavola, Sami. 2005. "Peircean Abduction: Instinct or Inference?" *Semiotica* 2005, no. 153: 131–154. https://doi.org/10.1515/semi.2005.2005.153-1-4.131.

Papetti, Randy, Paige Kaneb, and Lindsay Herf. 2019. "Outside the Echo Chamber: A Response to the Consensus Statement on Abusive Head Trauma in Infants and Young Children." *Santa Clara Law Review* 59, no. 2: 299–368. https://digitalcommons.law.scu.edu/lawreview/vol59/iss2/1.

Park, Woosuk. 2015. "On Classifying Abduction." *Journal of Applied Logic* 13, no. 3: 215–238. https://doi.org/10.1016/j.jal.2015.04.001.

Peirce, Charles Sanders. 1932. *Collected Papers of Charles Sanders Peirce, Volumes I and II: Principles of Philosophy and Elements of Logic*. Edited by Charles Hartshorne and Paul Weiss. Cambridge, MA: Harvard University Press. https://edisciplinas.usp.br/pluginfile.php/5165117/mod_resource/content/0/The%20Collected%20Papers%20of%20Charles%20Sanders%20Peirce%20%282904s%29.pdf.

———. 1935. *Collected Papers of Charles Sanders Peirce, Volumes V and VI: Pragmatism and Pragmaticism and Scientific Metaphysics*. Edited by Charles Hartshorne and Paul Weiss. Cambridge, MA: Harvard University Press. https://edisciplinas.usp.br/pluginfile.php/5165117/mod_resource/content/0/The%20Collected%20Papers%20of%20Charles%20Sanders%20Peirce%20%282904s%29.pdf.

———. 1958. *Collected Papers of Charles Sanders Peirce, Volumes VII and VIII: Science and Philosophy and Reviews, Correspondence and Bibliography*. Edited by Arthur W. Burks. Cambridge, MA: Harvard University Press. https://edisciplinas.usp.br/pluginfile.php/5165117/mod_resource/content/0/The%20Collected%20Papers%20of%20Charles%20Sanders%20Peirce%20%282904s%29.pdf.

———. 1992. *The Essential Peirce: Selected Philosophical Writings*. Edited by the Peirce Edition Project. Indianapolis: Indiana University Press.

Pietarinen, Ahti-Veikko and Francesco Bellucci. 2014. "New Light on Peirce's Conceptions of Retroduction, Deduction, and Scientific Reasoning." *International Studies in the Philosophy of Science* 28, no. 4: 353–373. https://doi.org/10.1080/02698595.2014.979667.

Plutynski, Anya. 2011. “Four Problems of Abduction: A Brief History.” *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 1, no. 2: 227–248. https://doi.org/10.1086/660746.

Price, Kelly J. and Kenna J. Miskelly. 2015. “Why Ask Why? Logical Fallacies in the Diagnosis of Fetal Alcohol Spectrum Disorder.” *Ethics & Behavior* 25, no. 5: 418–426. https://doi.org/10.1080/10508422.2014.946031.

Ramoni, M., M. Stefanelli, L. Magnani, and G. Barosi. 1992. “An Epistemological Framework for Medical Knowledge-Based Systems.” *IEEE Transactions on Systems, Man, and Cybernetics* 22, no. 6: 1361–1375. https://doi.org/10.1109/21.199462.

Reece, Robert M. 2001. “Chapter Eighteen: Controversies in Shaken Baby/Shaken Impact Syndrome.” *Journal of Aggression, Maltreatment & Trauma* 5, no. 1: 367–388. https://doi.org/10.1300/J146v05n01_18.

Robinson, Thaddeus. 2021. “Chapter 14: Inductive Generalizations.” In *Arguments in Context: An Introduction to Critical Thinking*, by Thaddeus Robinson. Allentown, PA: Muhlenberg College. https://open.muhlenberg.pub/arguments-in-context/chapter/inductive-generalizations/.

Roche, William and Elliott Sober. 2013. “Explanatoriness Is Evidentially Irrelevant, or Inference to the Best Explanation Meets Bayesian Confirmation Theory.” *Analysis* 73, no. 4: 659–668. https://doi.org/10.1093/analys/ant079.

Saunders, Dawn, Maria Raissaki, Sabah Servaes, Catherine Adamsbaum, Arabinda Kumar Choudhary, Joëlle Anne Moreno, Rick R. van Rijn, and Amaka C. Offiah. 2017. “Throwing the Baby out with the Bath Water: Response to the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU) Report on Traumatic Shaking.” *Pediatric Radiology* 47, no. 11: 1386–1389. https://doi.org/10.1007/s00247-017-3932-8.

Schupbach, Jonah N. 2017. “Inference to the Best Explanation, Cleaned Up and Made Respectable.” In *Best Explanations: New Essays on Inference to the Best Explanation,* edited by Kevin McCain and Ted Poston, 39–61. Oxford: Oxford University Press.

Schurz, Gerhard. 2016. “Common Cause Abduction: The Formation of Theoretical Concepts and Models in Science.” *Logic Journal of the IGPL* 24, no. 4: 494–509. https://doi.org/10.1093/jigpal/jzw029.

Sober, Elliott. 1988. *Reconstructing the Past: Parsimony, Evolution, and Inference*. Cambridge, MA: Massachusetts Institute of Technology Press.

Stanford, P. Kyle. 2006. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press.

Strouse, Peter J. 2016. “Child Abuse: We Have Problems.” *Pediatric Radiology* 46, no. 5: 587–590. https://doi.org/10.1007/s00247-016-3551-9.

———. 2018. “Shaken Baby Syndrome Is Real.” *Pediatric Radiology* 48, no. 8: 1043–1047. https://doi.org/10.1007/s00247-018-4158-0.

Tait, Gordon. 2009. “The Logic of ADHD: A Brief Review of Fallacious Reasoning.” *Studies in Philosophy and Education* 28, no. 3: 239–254. https://doi.org/10.1007/s11217-008-9114-2.

Thagard, Paul R. 1978. “The Best Explanation: Criteria for Theory Choice.” *Journal of Philosophy* 75, no. 2: 76–92. https://doi.org/10.2307/2025686.

Upshur, Ross. 1997. “Certainty, Probability and Abduction: Why We Should Look to C.S. Peirce rather than Gödel for a Theory of Clinical Reasoning.” *Journal of Evaluation in Clinical Practice* 3, no. 3: 201–206. https://doi.org/10.1046/j.1365-2753.1997.00004.x.

Van Fraassen, Bas C. 1989. *Laws and Symmetry*. Oxford: Oxford University Press.

Williamson, Timothy. 2018. *Doing Philosophy: From Common Curiosity to Logical Reasoning*. Oxford: Oxford University Press.

Woods, John. 2013. *Errors of Reasoning. Naturalizing the Logic of Inference*. London: College Publications.

Yu, Shiyang and Frank Zenker. 2018. “Peirce Knew Why Abduction Isn’t IBE: A Scheme and Critical Questions for Abductive Argument.” *Argumentation* 32, no. 4: 569–587. https://doi.org/10.1007/s10503-017-9443-9.

Zimmerman, Robert A., Larissa T. Bilaniuk, Derek Bruce, Luis Schut, Barbara Uzzell, and Herbert I. Goldberg. 1979. “Computed Tomography of Craniocerebral Injury in the Abused Child.” *Radiology* 130, no. 3: 687–690. https://doi.org/10.1148/130.3.687.