Synthese (2024) 203:194
https://doi.org/10.1007/s11229-024-04636-6

ORIGINAL RESEARCH

On informational injustice and epistemic exclusions

Abbas Bagwala[1]

Received: 4 April 2023 / Accepted: 15 May 2024 / Published online: 3 June 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

Information is a unique resource. Asymmetries that arise out of information access or processing capacities, therefore, enable a distinctive form of injustice. This paper builds a working conception of such injustice and explores it further. Let us call it informational injustice. Informational injustice is a consequence of informational asymmetries between at least two agents, which are deeply exacerbated due to modern information and communication technologies but do not necessarily originate with them. Informational injustice is the injustice of having information from an informational surplus being used to disadvantage the agent with less information.

This paper argues that informational injustice exploits an agent as a knower, specifically exploiting the agent's limitation in possessing or processing information—an agent is exploited because she is not informed or lacks in her ability to process information. In the case of lack of information, the agent simply does not know the information under consideration; a person is algorithmically manipulated or nudged to buy a product or vote for someone. In the case of a lack of capacity to process information, the agent simply cannot use the information, despite having access to it, to reach epistemically valuable states such as knowledge; a lawyer dupes you because he knows more about the inner workings of a courtroom and the law. Technically, you have access to the information your lawyer has, but you cannot make use of it due to constraints on time and cognitive effort. Informational injustice excludes the harmed agent from participating in knowledge practices. Thus, informational injustice is also a kind of epistemic exclusion.

After fixing the concept of informational injustice, the paper distinguishes between two kinds of informational injustices: interactional informational injustice and structural informational injustice. The former concerns interactions between agents, while the latter concerns social structures that emerge out of interactions between agents.

Keywords Information · Epistemic · Injustice · Data · AI · Technology

Extended author information available on the last page of the article

1 Introduction

Tribal populations, when governed by a state with a majority of non-tribal people, are often misled into thinking that the effect of certain laws on their lands would actually benefit them when they would not. Privatization of tribal lands for coal-mining, for instance, comes to mind. The use of force aside, one reason why people are misled is because they do not have enough information or the capacity to understand the information presented to them. Often, this is due to the fact that the state that governs the land has an official language of the law that a tribal population may simply not understand. Since conflicts over land use and tenure are often fought in the courts of the imposing state, the disadvantage of failing to understand information is a substantial one. The ignorance of laws and their implications makes solidarity within these communities hard to form, thus deterring them from opposing an idea that will ultimately disadvantage them. The communities are, to a great extent, disadvantaged because of a lack of information. This is an example of informational injustice. It happens when an informational asymmetry is used to disadvantage the less-informed agent(s). Some other examples include a homeopath falsely claiming that a placebo pill would work, often taking advantage of the patient's inability to understand the information on the subject; tobacco companies hiding studies that link the use of tobacco to cancer, and only presenting information in favor of the use of tobacco, taking advantage of the consumer's lack of information; a lawyer duping their client because she knows more about the inner workings of the courthouse, and so on. All these disadvantages are such that they would cease to exist if the person or groups being disadvantaged had the information the more-informed person or group does.

With these examples in mind, I aim do two things in this paper. Firstly, I show how precisely the informational advantages or asymmetries, like the ones shown above, lead to informational injustice. In doing so, I use Luciano Floridi's (2011) definition of information to form a working conception of informational injustice. Secondly, owing to the unique nature of *information* as a resource, I show how any instance of an informational injustice is necessarily an epistemic exclusion, that is, it excludes agents from participating in knowledge practices. Moreover, the harm associated with informational injustice manifests not only in terms of unequal distribution (e.g. the unjust distribution of rice or wood among people) but also as injustice done to people in their ability as knowers.[1] I argue that this distinction is crucial because it clarifies how mere redistribution of information cannot *always* solve ethical problems arising out of disadvantages caused to agents due to informational asymmetries. This is because human agents have limited informational capacities that mere redistribution of the resource of information cannot always mend. If we are to fix informational injustice, it would do well to take a structural approach, keeping in mind that agents have limited informational capacities, rather than to work only on reducing informational asymmetries by redistributing information. A structural approach

[1] One may contrast this to Fricker's use of the phrase for agents in their 'capacity as knowers.' The harm Fricker has in mind is prejudicial and related to power structures; hermeneutic injustice and testimonial injustice result because we refuse to listen to people due to interactional or structural prejudice. Informational injustice is not necessarily prejudicial. It is the willful use of an informational asymmetry to disadvantage other agents; therefore, it is an advantage taken of their ability, or the lack thereof, as knowers.

would ensure we stop ascribing blame on agents for being unable to continuously know more (which is the logical consequence of the untenable 'notice and consent' framework) but instead find global solutions that take varying informational capacities as given. Thus, one consequence of my approach is that we stop looking at the redistribution of information as key to solving issues arising out of informational asymmetries.

In Section 2, I define information, clarifying why I prefer using Floridi's modification of the general definition of information. I then define informational asymmetry and introduce the conceptual distinction between first-level and second-level information and informational surplus. Both concepts are used as heuristics to appreciate the novel nature of information as a resource.

In Section 3, I define informational injustice. Next, I show why and precisely how the harm it is concerned with is a form of epistemic exclusion.[2] I do this by contrasting informational injustice with Fricker's (2007) and Harris' (2023) use of the term 'epistemic' in their conceptions of injustices. Where Fricker (2007) requires a wrong be done to an agent in her capacity as a knower, Harris' *epistemic domination* (2023) requires only that an agent (at a lower position in an asymmetric epistemic relation) not be allowed to reach "epistemically valuable states." I show how Fricker's and Harris' seemingly different uses of the term have one thing in common, namely the idea of disallowing an agent to participate in knowledge practices. Since informational injustice is also a kind of harm that disallows agents from participating in knowledge practices, it is similar to epistemic injustice and epistemic domination. All three, I argue, are epistemic exclusions; that is, they are practices that disallow agents from participating in knowledge practices either by keeping information from them, burdening them with too much information, or by not listening to their testimony or experiences.

2 Information and informational asymmetry

2.1 What is information?

By information, I mean merely semantic information. In his treatment of information, Floridi (2011) proposes we modify the general definition of information. Under what Floridi considers to be the general definition, semantic information is information that (a) consists of n data such that $n \geq 1$, where (b) the data are *well-formed*, and (c) the data are *meaningful.* In other words, any data that are meaningful and interpretable may constitute information. The general definition of information is accepted as a reliable definition of information (Floridi, 2011). What interests us in the general definition is the possibility for information to be classified as false. One notes that under such an understanding of information even false propositions constitute information. This, Floridi maintains, is an implausible and insufficient understanding of

[2] That is, informational injustice is necessarily an 'epistemic' harm because it disallows agents from participating in knowledge practices. It is not, however, a form of 'epistemic injustice' in Fricker's (2007; 2011) sense of the term. More on this later.

semantic information because it does not encapsulate truth. Why?[3] One reason for this is that this definition of semantic information does not allow for the semantic destruction of information. But what is the semantic destruction of information and why is it important to us?

Consider the following true proposition, P1: "Russell was born in Wales, long after Newton had passed." Imagine this is stored on a piece of paper you have. The content of the information will be lost if your dog eats the paper—this is the physical loss of information. You may also lose the content of the information if the handwriting is generally bad, where it becomes hard to understand what is written on the paper—this is the syntactical loss of information. Now consider that you have another proposition, P2: "Russell was born in Wales, long after Einstein had passed," added to the same sheet of paper. If considered together, under the general definition of information where information does not encapsulate truth, P1 and P2 both constitute information. Therefore, information in P1 + P2 > information in P1; PI + P2 together have more information than only P1.

However, consider now P3: "It is not true that (Russell was born in Wales, long after Newton had passed)." In other words, P3 = ~P1. Let us assume P3 is added to our sheet of paper instead of P2. Now, we have two propositions on the paper that contradict each other. We know that an agent in possession of P1 and an agent in possession of P3 would both have information under the general definition. However, *ceteris paribus* (that is, given that there is no syntactical or physical loss of information), an agent in possession of P1 and P3 would also have information, if not more than when considered separately.

But this raises a core problem for us. We are forced to admit that agents with more data[4] are necessarily more likely to be agents in possession of more information. That is, data that contradict each other do not cancel out. On the other hand, if we understand information as only true and meaningful data, then we may be allowed to conclude that being merely in possession of more data does not necessarily mean one is in possession of more information; a mere surplus of data does not *de facto* constitute an informational surplus. This is an important distinction because we wish to characterize a unique injustice that is based on keeping agents from knowledge about the world. Our concern is the harm done to agents because of a lack of availability of facts or their inability to process too many facts. In other words, we wish to acknowledge and explore the harm that can be done to agents either because they do not know the truth or cannot know the truth.

Finally, we have more than taxonomic reasons to choose Floridi's definition of information. As Floridi has made sufficiently clear (2011), the modified definition of information, one that allows only data that is true to be classified as information, also helps solve the Bar-Hillel–Carnap Paradox[5] in information theory. We thus adopt the modified general definition of information. Information becomes, then, well-formed

[3] The interested reader must refer to Floridi (2011) for a full-fledged account Floridi's work on the general definition of information, leading to his theory of strongly semantic information.

[4] Henceforth, 'data' will be used interchangeably with 'well-formed and meaningful data.'

[5] The paradox: "A self-contradictory sentence, hence one which no ideal receiver would accept, is regarded as carrying with it the most inclusive information." Cited in Floridi, 2011, p. 100.

and meaningful data that is also true. Note that false information now becomes an oxymoron (Floridi, 2011).

2.2 What is an informational asymmetry?

Informational asymmetry is caused due to a difference in the number of informative propositions that two agents hold. Informational asymmetry, in other words, is the condition in which there is a non-equivalence of information between two (or more) agents. Understood as such, informational asymmetries are quite common and exist between any two agents. You will always have information about something I do not, and I will always have information about something you do not. So far, so good. A difference in the amount of information held by two agents is not *ipso facto* a cause of concern.

Some informational asymmetries are the result of natural processes; a human adult would possess information that a human baby would not, or a difference in cognitive capacities between normal and differently-abled human adults would result in a difference in their capacities to possess and process information, resulting in both holding different amounts or kinds of information. Some informational asymmetries are the result of socio-political activity; a human adult in possession of a computer or with access to a huge library has the access and the potential to store more information than only a normal human adult. It is the latter (i.e. informational asymmetries caused due to socio-political activity) more than the former that concerns us here.[6]

In its simplest form, an informational asymmetry would be when the information one agent holds is the subset of the other.[7] However, before we proceed to define it, the 'holding' of information needs some explanation. What does it mean for an agent to 'hold' information? An agent may be said to be holding information if he or she is in possession of the information. We may make a distinction between an agent having 'access' to information and an agent 'holding' information. An agent has 'access' to information when he or she can find out the information from a source but has not yet found out the information. Consider any privacy statement for a software application you have on your cellphone, the one you accept without reading fully. You can always go back to it and get the information you require, but you do not 'hold' it yet.

[6] However, the former also raises interesting moral or ethical considerations between two parties with differing abilities to treat information. Consider, for instance, how a young child may be fooled into believing certain things about the world, curated such that they may do something to one's advantage. Indoctrinating a child certainly forms an interesting case of informational injustice.

[7] I thank my first reviewer for pointing out a potential confusion here. The unification of qualitative (meaning-making semantic/agent-based) and quantitative (Shannon information or Kolmogorov complexity) theories is an open problem in the philosophy of information. The current paper does not deal with that issue. My treatment of information is purely qualitative (in the sense given here), where the meaning-making of a proposition done by an agent is taken for granted. Since meaning in propositions is assumed, how quantitative information becomes qualitative should not have a bearing on the results of my paper. If the reader is unconvinced with this line of reasoning, he or she may simply substitute the word 'information' with 'semantic information' throughout the paper, assuming that whenever we speak of transfer of information, it is between agents capable of understanding the meaning of that information.

Holding information, on the other hand, is to be in possession of information such that one may transmit it (e.g. utter it to oneself or others).[8]

Thus, we may define informational asymmetry like so:

> An informational asymmetry is a condition between two (information processing) agents such that there is at least one topic 'S' about which one of the agents holds more informative propositions than the other.

This is the minimal condition under which an informational asymmetry persists. Generally, informational asymmetries are far more complex. Between two agents, one may know some things about several topics, and the other may know some other things about several topics, half of which the latter may share with the former in a complex web such that there are topics where one has more information about a few things and the other about a few others. Given that the two agents are humans (let us call them Alice and Burhan), we know that they may also formulate new information from what they already possess. Suppose regarding the chess player Mikhail Tal, both agents have the information that (a) he was from Riga. Alice, however, also knows that (b) Riga is a part of Latvia. Alice may then conclude, inferring by deduction, that (c) Mikhail Tal was also Latvian.

2.3 First-level and second-level surplus information

In our case, Alice has more information than Burhan about the topic 'S.' This is *surplus information.* From the initial information that she had over Burhan, Alice has now managed to increase this surplus. Let us call the information both Alice and Burhan originally had first-level information. The increase in information due to information *inferred* from the information that the agents already had, we may now call second-level information or inferred information.[9] Consequently, the surplus information that Alice had over Burhan may now be called *first-level surplus information*, and the information Alice derived from the information she already possesses, such that Burhan does not hold the same, may be called *second-level surplus information.*[10]

A note of caution is in order here. The idea that information inferred from preexisting information due to deduction is *new* information is a substantially controversial idea, a position clearly rejected by the logical empiricists of the Vienna Circle,

[8] Arguably, for human agents, the 'holding' of propositions implies adding the words to a network of already present words. A human agent's holding of information is fundamentally different from a computer's because the latter may be able to hold information without cumulatively adding it to the larger whole, i.e. in one corner of a hard drive that works merely on changes in transistor states.

[9] Note that inferring information requires an agent that understands the meaning of semantic information. In my case, this is only humans. There is debate about whether computers understand semantic content, including some claiming that semantic understanding emerges at a certain hyperparameter in LLMs. I seriously doubt this, but find it reasonable to not foreclose the possibility (at least for some notion of 'understanding').

[10] This is not an ontological claim. I use the notion of first-level and second-level information as heuristics to appreciate the idea that differing computational or cognitive capacities may result in differing abilities to make inferences. I thank my third reviewer for pointing out a potential confusion arising here.

who considered the truths of logic and mathematics as necessary and independent of experience. According to Hempel, for instance, a purely logical deduction simply "re-asserts part of what has already been stated in the premises" (Hempel, 1945). For Quine, the truths of logic are analytic in the "traditional sense of the word, that is to say by virtue of the meaning of the words" (Bergström & Føllesdal, 1994).

In contrast, D'Agostino and Floridi (2009) argue that classical deduction, even at the propositional level, is actually informative. It is not in our interest to go into the details of their arguments here. The present author is inclined towards the latter view because it resolves the "scandal of deduction" (Hintikka, 1973), not rendering much or all of logic and mathematics such that no new discovery in them is considered 'informative.' The latter view also resolves difficulties arising with both treating methods of inferences such as deduction as some form of intellectual psychoanalysis (Hempel, 1945) or, per Wittgenstein, some "imperfection in our logical language" (D'Agostino, 2013).

The reader inclined to the former view may be less inclined to accept the view that Alice's conclusion that Mikhail Tal is Latvian gives Alice any *more* information than that residing in Alice when she is holding the information that Mikhail Tal was from Riga and that Riga is a part of Latvia. In other words, such a reader may not agree that Alice's deduction of (c) results in Alice's net increase in information. However, insofar as one grants that given an informational surplus between two agents with equal computational resources or cognitive capacities, the agent with more information certainly has an equal if not higher capacity to make more inferences such that she may *know* the world better, my arguments will still hold. Here, the reader may again appreciate the importance of having defined information as well-formed and meaningful data that is also true; merely a surplus of well-formed and meaningful data cannot lead to such an advantage.

Now, back to Alice and Burhan. To make our analysis easier, we may assume—without any loss of understanding—that the information Burhan holds is a subset of the information Alice holds. The extra information Alice possesses may be called *surplus information.* Imagine Alice is in possession of a powerful computer that allows her to process large datasets, such that Alice's computational prowess or informational capacity[11] is increased. Consider now that part of the surplus information Alice holds is a dataset on the commute habits of all the people residing in a neighborhood consisting of 10,000 people. Suppose that Alice models the commute behavior of the 10,000 people and derives insights using her computational prowess. Thus, from the additional information that Alice already had, she was able to formulate more information, increasing her surplus information. From Alice's perspective, we may call the initial surplus information (the dataset with the commute habits of 10,000 people) *first-level surplus information*, and the information she later derived from that (the results due to derivation of insights) *second-level surplus information*.

This allows us to acknowledge a trend the rise of powerful information and communication technologies has caused. Agents with advanced informational capacities

[11] Informational capacity, or computational prowess, is simply one's capacity to make a given number of reasonable inferences in a given amount of time. More inferences per unit of time imply more informational capacity.

(e.g. Alice with access to supercomputers) may be able to generate more information than agents with lesser informational capacities (e.g. Burhan with access to his personal computer), with everything else remaining the same. Thus, the amount of second-level information in an information surplus *seems to be* directly proportional to the difference in the informational capacities between two agents. Alice's informational surplus, given the same amount of first-level information, seems to increase in direct proportion to Alice's computational or cognitive prowess over Burhan. The more information Alice can infer from the first-level information she has, the more her advantage over Burhan is, informationally speaking. But, the presence of an asymmetry is no cause for concern unless it is used to disadvantage the less-informed agent. To understand that is the purpose of this paper, and we explore it next.

3 Informational injustice

3.1 What is informational injustice?

We are now in a position to formulate a working conception of informational injustice. Informational injustice occurs when a more-informed agent disadvantages a less-informed using the surplus information the former has. Informational injustice occurs if and only if an agent is wronged owing to his or her lower position in an informational asymmetry. We can resolve a persisting informational injustice only by ensuring that either the disadvantage ceases to exist, the informational surplus ceases to exist, or both.

But what sort of disadvantage do I mean, precisely? Consider the case of a lawyer duping his clients or the doctor duping his patients again. What sort of disadvantage do the less-informed parties face? In the lawyer's case, the client might have to end up paying a lot more money than she actually needs to, and she might have to dedicate a lot more time than required. Both these things—and anyone who has experienced the matters of the courtroom closely may testify—may exhaust the client emotionally sufficiently enough to feel drained. The disadvantage, therefore, can easily be life-altering. Thus, it is a kind of material disadvantage in a very real sense. The case of the doctor is similar. When people find themselves in medical situations, they are often very prone to making anxious decisions, and the trust they bestow on their doctor is very crucial. To be disadvantaged when in a position such as that may cause them substantial suffering. The only way for the less-informed agent to *not* be disadvantaged in both these cases is for him or her to first get hold of the information from the informational surplus between him or her and the more-informed agent. In other words, to get rid of the informational surplus.

3.2 Informational Injustice vs. 'epistemic' injustices

Informational injustice is an epistemic exclusion in the sense that it disallows agents from participating in knowledge practices, much like epistemic injustice (Fricker, 2007) and epistemic domination (Harris, 2023). Informational injustice, epistemic injustice, and epistemic domination are all kinds of epistemic exclusions. Epistemic

exclusions are practices that disallow agents from being participants in knowledge practices.[12] Information injustice is not a kind of *epistemic* injustice in Fricker's (2007; 2011) sense of the term. However, it is similar to epistemic injustice because it is a harm that does not follow the distributive justice logic. That is, a simple redistribution of information would not solve problems caused by informational advantages.

Since the word *epistemic* is widely used in contemporary literature in social epistemology in a variety of ways (Fricker, 2007, 2011; Dotson, 2011; Medina, 2011; Pohlhaus, 2012; Origgi, 2012; McKinnon, 2016; Symons & Alvarado, 2022; Harris, 2023), some clarification of its use is demanded. Miranda Fricker's (2007, 2011) use of the term is most prominent in the literature on epistemic harms. In Fricker's use of the term, epistemic injustice necessarily satisfies broadly two conditions: Firstly, it is harm done due to a resource that does *not* follow the distributive justice logic (Fricker, 2007). Secondly, the injustice harms an agent in his or her *capacity as a knower* (Fricker, 2007). Fricker's aim is to:

"… characterize two forms of epistemic injustice: testimonial injustice, in which someone is wronged in their *capacity* as a giver of knowledge; and hermeneutical injustice, in which someone is wronged in their *capacity* as a subject of social understanding." [Italics mine].

In Fricker's understanding of epistemic injustice, the socio-political standing of the agents involved is necessary for the injustice to be exercised. In other words, the socio-political standing of an agent determines how any given epistemic injustice will manifest in a given situation. Thus, the injustice cannot be understood when abstracted from its socio-political context.

Contrast this to Harris' 'epistemic domination,' which is understood as "the non-mutual capacity of one party to control the evidence available to another" such that "the dominated party may have difficulty attaining epistemically valuable states" (Harris, 2023). In Harris' use of *epistemic*, the harm done concerns only the intended control of evidence by one party to the disadvantage of the other such that the harmed party does not reach 'epistemically valuable states.' For Harris, these epistemically valuable states refer to acquiring knowledge, justified belief, or understanding. Here, the socio-political standing of the agents involved is not central to the harm under consideration. In other words, it does not matter *who* the agents involved are; the harm sufficiently occurs just by virtue of the fact that one party did not allow the other party to attain epistemically valuable states. Epistemic domination, therefore, can be abstracted from the socio-political context in which it occurs.

Fricker's and Harris' use of the word *epistemic* for harms relating to knowledge clearly have stark differences. However, what underlies both and informational injustice is the idea of *disallowing agents from participation in knowledge practices.* In Fricker's epistemic injustice, agents are not allowed space, due to prejudice (interactional or structural), to participate in knowledge practices (in comprehending or producing knowledge). In Harris' epistemic domination, the dominated agent is dom-

[12] Not all practices should rationally involve all agents and not all epistemic exclusions are wrong. For instance, it would make little sense to have sociologists weigh in on quantum entanglement and toddlers weigh in on what path to choose on a dangerous trek.

inated precisely because she is not allowed evidence. By not being allowed to be in possession of evidence, she, too, is not allowed to participate in knowledge practices.

With this in mind, consider now informational injustice. Suppose there is an informational asymmetry between two agents. There are two ways in which the less-informed agent may be harmed informationally. The more-informed agent may disadvantage the less-informed agent simply because the latter does not have some information the former does, e.g., by *hiding* information from the latter. Epistemic domination is one example of this.[13] Note that the act of hiding the information is not, by itself, an informational injustice. For the informational asymmetry to advance to informational injustice, the more-informed agent must also use the information surplus to disadvantage the less-informed agent. This is a trivial case of informational injustice, where simply having more information may often put someone who has the means to disadvantage another in the position to do so.

It might seem that because informational injustice is a matter of disadvantaging an agent when they do not have information, redistributing the information to the less-informed agent may prevent him or her from being harmed. But this is precisely where our problem starts. For human agents, pieces of information (that is, of well-formed, true, and meaningful propositions) about the world do not stand in isolation from each other but form a coherent whole—knowledge is a network of propositions. It takes cognitive effort and time to form such a network, a coherent whole where meaningful associations between propositions (a semantic understanding) are achieved.

Therefore, the more-informed agent may also disadvantage the latter by taking advantage of the latter's limited information processing or cognitive capacities. In this case, information may be available to the less-informed agent if she seeks to access it. However, given the real constraints of time and cognitive effort, the agent is simply incapable of doing so. The more-informed agent harms the less-informed agent epistemically by taking advantage of her limited information processing or cognitive capacities that disallow the agent to reach epistemically valuable states. In the former case (when information is simply unavailable), the problem that causes the harm is a lack of information. In the latter case, the problem that causes the harm is a lack of information processing or cognitive capacities.

In either case, the less-informed agent cannot be a participant in knowledge practices. When information is simply lacking with the agent, she cannot be part of a knowledge practice because information—the raw material of knowledge—is missing. However, as I remarked earlier, this is somewhat of a trivial case. But what happens when information is indeed available with the less-informed agent (remember, the agent is not yet *holding* that information,[14] she merely has access to it, say in a computer or a data center or a phone), but cannot process or use it? Can the agent still not be a participant in knowledge practices and thus be harmed? Answering these questions requires one final detour.

[13] Epistemic Domination (Harris, 2023) does not exhaust all possibilities.

[14] For a human agent, to hold information is to be able to answer a number of yes/no (Boolean) questions about a specific topic correctly (see above). If the human agent fails to answer yes/no questions about a given topic correctly, then he or she does not hold information about that topic. More on this below.

3.3 Floridi and knowledge as accounted information

Suppose you visit your lawyer, who is an excellent scholar on land ownership laws. Placed on his wooden table is a daunting yet comprehensive account of the information available on land laws, the kind of information that concerns you at the moment and can help you tremendously in building your case to sue your business partner, who has taken your fair share of the land away from you. You are paying your lawyer precisely because he knows the contents of the book, and you do not. However, since the copy of the book is right in front of you, its contents are available to you as well. The resource (information in the book) is redistributed and free to use by both parties. While you ponder about reading the book, wondering whether you should have chosen to study law instead of joining your dad's business, your cousin (a law student and a half-baked lawyer) walks into the room. There are now three people in the room whose knowledge of land ownership laws differs from each other. You know nothing about land ownership laws; your cousin can pass an exam on South Asian land ownership laws; your lawyer knows just about everything about land ownership laws. There is one artifact in the room that contains all the information (the book).

Intuitively, we may assert that the lawyer, you, and your cousin have differing amounts of knowledge on the subject of land ownership laws. We are now faced with a challenge: To give an account of why precisely the knowledge all three agents have on the subject of land ownership laws varies. Why do we say they have differing amounts of knowledge? Given that all three have access to all available information on the subject (the book), *how* do they know in differing amounts? In other words, even though all of them have access[15] to the same information, their knowledge of the subject of land ownership laws varies from each other's. In what sense can this be understood?

Fortunately for us, Floridi (2012) presents a network theory of account that formalizes our intuition above, namely that the three agents have differing amounts of knowledge, functionally speaking. The central question Floridi considers is this: How can (semantic) information be upgraded to knowledge? A brief sketch of his account below will help us finally solve the last piece of the puzzle, which—I must remind the reader—is the question: How can an agent, despite having *access* to information, still not be a participant in knowledge practices?

3.4 How can an agent, despite having access to information, still not be a participant in knowledge practices?

A small note before we move on. For those who wish to make no ontological commitments to the account of knowledge offered below, we consider the network account of knowledge to simply be a heuristic that proposes to answer the non-deviant "How do I know you have knowledge?" rather than the deviant "How do I know I have knowledge?" The difference is that the former allows for a functionalist account of knowledge that does not necessarily require doxastic commitments. This is not to

[15] Note of caution: they have 'access' to this information, but they do not 'hold' the information yet. See Section I, B.

say that knowledge does not require beliefs, of course, but merely that my best bet to know whether you have knowledge of 'x' is to ensure you are functionally able to fulfill any criterion of knowledge a given account offers. Insofar as the network theory of account is concerned, it is your ability to answer questions about 'x' that account for it. You may, of course, have knowledge of 'x' even if you cannot account for 'x' to me, but then I simply have no way of ensuring you do have knowledge of 'x.'

Floridi proposes we consider a graded conception of knowledge, where knowledge is upgraded information.[16] But how does information 'upgrade' to knowledge? Information upgrades to knowledge "if and only if it is correctly accounted for" (Floridi, 2012). There are two qualifications that a piece of information ought to make to be considered knowledge. Firstly, it should answer any "how come" questions it raises. Secondly, the questions it raises should be accounted for.

Floridi (2012) considers the question: How come the water in the electric kettle is boiling? To give an account of the information encoded in the question (that the water in the electric kettle is boiling), one needs to answer a number of "how come" questions:

> "How come" questions… may therefore receive different answers. "How come that the water in the electric kettle is boiling?" may receive as an answer "because Mary would like some tea" (teleological account), or "because Mary filled it with water and turned it on" (genealogical account), or "because electricity is still flowing through the element inside the kettle, resistance to the electric flow is causing heat, and the steam has not yet heated up the bimetallic strip that breaks the circuit" (functional account) (Floridi, 2012).

For the sake of simplicity, we ensure we convert any 'how come' questions into their Boolean forms. "For example: "how come that the water in the electric kettle is boiling?" may become "Is the water in the electric kettle boiling because Mary wants some tea?"" (Floridi, 2012). We may now consider a threshold of questions that an agent has to be able to answer about any given question, such that we may say the agent has knowledge about the phenomenon underlying the given question. For instance, to answer "How come that the water in the kettle is boiling?", we may require our model to answer, say, five related questions on the subject. Suppose you wish to understand whether your child has the scientific knowledge to answer the question about the boiling kettle. You may present your child with five Boolean[17] questions that, if she manages to answer, ensure that she knows about the boiling kettle if answered correctly ("Is heat being transferred to the water by radiation?" "Are electromagnetic waves required to heat the water?" "Is electricity required for the kettle?" etc.).

We may, thus, differentiate between merely having access to semantic information and possessing knowledge. You may give your child all the information on the

[16] I must remind the reader that information is well-formed, meaningful data that is also *true.* See the section of 'What is information?'

[17] The questions do not have to be Boolean; they can be non-Boolean. Boolean questions simplify the model.

subject, but if she fails to answer 'how come' questions arising out of a given piece of information, she does not have knowledge of the subject.

> There is a difference between semantic information and knowledge, which can be highlighted by epistemic luck[18]. The difference is that semantic information lacks the necessary structure of relations that allow different packets of information to account for each other. It follows that, for semantic information to be upgraded to knowledge, it is necessary to embed it in a network of relevant questions and corresponding correct answers (Floridi, 2012).

Now, we are in a good position to go back to the case of the lawyer. The lawyer, your cousin, and you have differing amounts of knowledge on the subject of land ownership laws because all of you can answer correctly a differing number of questions on the subject. For the question, "How come one requires a registered sale deed while selling land?" your lawyer will have more answers to account for the 'how come' questions arising out of this question than your cousin, who in turn has more answers than you do.

This gives us a graded conception of knowledge, making the intuition that about a given topic that some people know more than others, who in turn know even more, formalizable.[19] Some people know more than others because they can answer (to others or to themselves) more 'how come' questions about a given piece of semantic information correctly. To be part of a knowledge practice about a given topic (say land ownership laws), an agent must, therefore, be either able to comprehend or produce knowledge on the subject.[20] Without the ability to answer 'how come' questions, this is impossible.

Thus, answering the question that motivates this section: An agent, despite having access to information, can still *not* be a participant in knowledge practices (thus making the harm associated with the phenomenon 'epistemic') if he or she is unable to answer 'how come' questions about the given piece of information. The ability to answer 'how come' questions ensures that the agent both (a) holds the information and it is not due to (b) epistemic luck, thus ensuring that the agent has knowledge.

3.5 Back to informational injustice: what sort of disadvantage are we talking about? What injustice?

I defined informational injustice as the disadvantage a more-informed agent does to a less-informed agent due to the informational asymmetry between the two. I showed how this manifests in two ways depending on whether the less-informed agent has access to information or not. In the case where the less-informed agent simply does

[18] Epistemic luck is simply finding oneself in possession of information (well-formed, meaningful, and true data).

[19] The interested reader should refer to Floridi (2012) for a full-blown account of the network theory of account and semantic information.

[20] In the case of Fricker's testimonial injustice and hermeneutic injustice, the agent is disallowed from participating due to prejudice, structural or otherwise.

not even have access to the surplus information, the more-informed agent simply uses the surplus information to disadvantage the less-informed agent. But a bit more needs to be said about the case where the less-informed agent does have access to information but can still be disadvantaged.

This is because we find ourselves in the latter situation more often than not. The internet enables tremendous access to information to agents, including open source repositories containing potential information[21] about how algorithms may even manipulate agents. The reason agents are disadvantaged is because it takes cognitive effort and time to convert that information you have access to into knowledge to be able to answer 'how come' questions about the information that is available.

Your visit to the lawyer is simply your way of not having to take the cognitive effort in trying to answer 'how come' questions related to sale deeds, for instance.[22] Your visit to the doctor to know your ailment, the professor for quick guidance regarding the next steps you should take, or the interior designer to decorate your house are some more examples of you choosing to take advice for something than putting the cognitive effort in knowing that thing fully yourself (or answering 'how come' questions about it, e.g. 'how come my nose is still running?' 'how come I should not read Kant now?' 'how come yellow looks better than black here?' etc.).

But what sort of a disadvantage am I talking about precisely? Earlier, I gave a few examples involving a lawyer duping his client or a doctor duping her patients. The source of the injustice done in these cases is epistemic since all such cases involve taking advantage of the less-informed agent's lack of information or his/her inability to hold some information (such that holding that information would have enabled them to know about the subject under consideration). However, the resultant harm is not necessarily epistemic.[23] The harm done to the less-informed agent may be material in a very real sense. If the less-informed patient is made to spend money on unnecessary tests, he may lose a lot of money that he may have urgently required elsewhere, for instance.

Informational injustice can only be understood positively, for it manifests only when a disadvantage is done actively to a less-informed agent. Much like Fricker's (2007; 2011) testimonial and hermeneutic injustice, only the active presence of agents' disadvantage can shed light upon informational injustice.

Consider targeted political advertising, which is the phenomenon where software applications, such as social media applications on your phone, advertise only particular political advertisements on your feed, tailored specifically to you based on your online behavior. Targeted political advertising is often advantageous because it helps you find useful and interesting pieces of information for you out of a sea of information on the internet. However, it often manifests harmfully, such as when it is used "to influence… individuals' decision-making by appealing to subconscious mechanisms

[21] Potential information and not 'information' because data available is not always well-formed and meaningful but needs to be converted first to be so. I thank my first reviewer for pointing this out.

[22] This is not to deny, of course, the non-cognitive work the lawyer's office also does for you, including the clerical legwork.

[23] It can be, as in the case of Harris' epistemic domination (2023), which we may now consider as one case of informational injustice.

or known irrational proclivities" (Saetra, 2021). In such cases, a political campaign may take advantage of your lack of information and bombard you with only specific information that your online behavior has shown the most proclivity to believe. This is an example of informational injustice because the algorithms enable the agents creating those software applications to know things about you that you cannot actively keep track of, like subconscious mechanisms and irrational proclivities.

Consider, now, the long 'terms of use' statements that similar software applications ask you to read before signing up for them. Maliciously written statements allowing the software applications to access your phones in ways that you would find hard to foresee if you did not know the whole statement is another example of informational injustice. Very few people, if any, have the time and cognitive effort required to read and know the contents of these long statements, which are often written in a convoluted language by design.

Finally, one may also consider government policies that certain sections of a population do not understand because of language barriers, where a government's official language differs from the language of, say, a given tribal population. When the tribal population is unable to understand how the laws of the land affect them precisely, they may be at the receiving end of a disadvantage due to a lack of information, especially when they must fight legal battles to assert their rights.

Keeping in mind the kinds of disadvantages above, we understand the disadvantages caused toagents to be both epistemic and material. How precisely an agent's life is affected cannot be answered in the abstract, but only understood on a case-by-case basis. All examples above treat the less-informed agents unfairly. The unfairness is caused due to either an agent's lack of information, in which case the agent does not even know that he or she does not know something, or due to the agent's inability to process information, in which case the agent—often not due to lack of trying—cannot know. The unfairness in either case is unjust because it is enabled only by the less-informed agent's being in a position of lesser informational power.

3.6 Biting a bullet

But while informational injustice is always an epistemic exclusion, what do we say about cases where it is required in order to fight an arguably greater injustice? What can we say about cases where a workers' union decides to go up against its bosses?[24] After all, the union might have to hide some information from its bosses to be able to coordinate the actions of the union members. Clearly, the union would, therefore, be committing an informational injustice because they would use an informational asymmetry to disadvantage their bosses and, say, eventually making them materially less well-off. In such a case, I argue, it would be appropriate to bite the bullet. This should have no bearing on the conceptualization of informational injustice. In going against the bosses for a greater good, the union members may also have to commit to other unfair or unjust practices, such as lying and manipulation. One may commit informational injustice and may still be working towards a greater good.

[24] I would like to thank my second anonymous reviewer for pointing this out as a potential conflict. It is also because of the second reviewer's comments that I wrote this section.

So, while informational injustice definitely excludes someone epistemically, whether or not it harms someone remains neutral unless applied under a moral framework. For instance, whether or not a parent in committing informational injustice against his child[25] is doing the right thing depends on the moral framework one operates under, which is quite similar to whether the parent's lying to the child is good or bad depends on whether one has deontological or consequentialist commitments.

3.7 Structural vs. interactional informational injustice

With the given examples, we now reach a final distinction that can be made based on the two ways in which informational injustice itself manifests. We may distinguish between interactional and structural critiques (Benn & Lazar, 2022).[26] The former concerns interactions between agents, while the latter concerns structures emerging out of interactions between agents. Consequently, 'interactional' informational injustice concerns the disadvantage one agent may put another agent at because of an informational surplus—like a doctor duping their patient. 'Structural' informational injustice concerns cases where informational asymmetries cause a structure (or a system) itself to become worse off owing to the misuse of surplus information by more-informed agents in the system—like the democratic *ethos* of democracy being lost because some people are capable of wielding enormously more information than others, causing disruption in some people's ability to choose freely. One may note that, even in structural informational injustice, because all agents are disadvantaged, less-informed agents will be disadvantaged *a fortiori*.[27]

4 Concluding remarks

I set out in this paper to develop a working conception of informational injustice—an injustice that is caused because of informational asymmetries between agents. I defined information, following Floridi (2011), as well-formed, meaningful data that is also true, arguing why it is a superior approach to defining information. Next, I defined an informational asymmetry—a condition under which one agent holds more informative propositions than the other. Finally, I defined informational injustice as the disadvantage a more-informed agent causes a less-informed agent *because of* or *by using* the surplus information the former has. I argued that informational injustice is an epistemic exclusion because it disallows the less-informed agent from participating in knowledge practices, much like epistemic injustice or epistemic domination.

Informational injustice is a wrong difficult to right not only because humans are agents with limited informational capacities but also because humans have access to

[25] By, say, disadvantaging the child by keeping her from enjoying something in the short-run based on an informational asymmetry; something that is good for the child in the long-run, but the child does not know that.

[26] Benn and Lazar (2022) use this conceptual separation in their paper on automated influence more generally to talk about privacy, among other things. We may use it for informational injustice as well.

[27] A full-blown account of structural informational injustice certainly demands another paper in its own right.

differing amounts of information and information processing technologies. An engineer with access to a data center may find it much easier to draw knowledge from large datasets than a school teacher in a village. The risk of informational injustice will remain so long as agents are in possession of different amounts of information or information processing technologies, both of which are inevitable in an information age. It would, therefore, do well to design systems that take for granted that participating agents in that system have, and will continue to do so, tremendously varying informational capacities. Such an approach is akin to the approach that demands a university make space for students with all sorts of capacities, including differently-abled students, who—for instance—may require ramps instead of stairs.

Acknowledgements Were it not for Ben Winokur, who very charitably called my initial paper 'something spicy,' I would have never sent it to publication. I am especially indebted to Saujanya Bharadwaj and Mahi Narayanan, whose insightful comments made the paper significantly more polished. I also thank Ramón Alvarado, Colin Koopman, Tammo Lossau, and Danny Weltman for their invaluable comments. Finally, my sincerest thanks to my reviewers. Their crucial remarks have made the paper much stronger than it initially was.

Declarations

- The author has no relevant financial or non-financial interests to disclose.
- The author has no competing interests to declare that are relevant to the content of this article.
- The author certifies that he has no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.
- The authors have no financial or proprietary interests in any material discussed in this article.

References

Benn, C., & Lazar, S. (2022). What's wrong with automated influence. *Canadian Journal of Philosophy*, 1–24.

Bergström, L., & Føllesdal, D. (1994). Interview with Willard Van Orman Quine in November 1993. *Theoria*, *60*(3), 193–206.

D'Agostino, M. (2013). Semantic information and the trivialization of logic: Floridi on the scandal of deduction. *Information*, *4*(1), 33–59.

D'Agostino, M., & Floridi, L. (2009). The enduring scandal of deduction. *Synthese*, *167*(2), 271–315.

Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia*, *26*(2), 236–257.

Floridi, L. (2011). *The philosophy of information*. OUP Oxford.

Floridi, L. (2012). Semantic information and the network theory of account. *Synthese*, *184*(3), 431–454.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

Harris, K. R. (2023). Epistemic Domination. *Thought: A Journal of Philosophy*.

Hempel, C. G. (1945). Geometry and empirical science. *The American Mathematical Monthly*, *52*(1), 7–17.

Hintikka, J. (1973). Logic, language-games and information, kantian themes in the philosophy of logic. *Revue Philosophique de la France Et de l*, *163*.

ignorance. *Hypatia*, *27*(4), 715–735.

McKinnon, R. (2016). Epistemic injustice. *Philosophy Compass*, *11*(8), 437–446.

Medina, J. (2011). The relevance of credibility excess in a proportional view of epistemic injustice: Differential epistemic authority and the social imaginary. *Social Epistemology*, *25*(1), 15–35.

Origgi, G. (2012). Epistemic injustice and epistemic trust. *Social Epistemology*, *26*(2), 221–235.

Pohlhaus, G. (2012). *Relational knowing and epistemic injustice*. Toward a theory of willful hermeneutical.

Saetra, H. S. (2021). *Big Data's Threat to Liberty: Surveillance, Nudging, and the Curation of Information*. Academic.

Symons, J., & Alvarado, R. (2022). Epistemic injustice and data science technologies. *Synthese*, *200*(2), 87.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Abbas Bagwala[1]

✉ Abbas Bagwala
abbasqb@protonmail.com; abagwala@uoregon.edu

[1] University of Oregon, Oregon, USA