

The Psychology of Pollution: Analysing How Heuristics Affect Perception of Air Quality and Policy Design

Introduction

Air pollution stands as one of the most pervasive and severe environmental causes of the 21st century, imposing a staggering toll on global public health and economic stability (WHO, 2016). Millions of premature deaths annually are attributed to exposure (WHO, 2016). Economic costs amounting to trillions of dollars—eroding national GDPs and straining healthcare systems—are starkly documented. Yet, societal and individual responses to this singular threat are not dictated solely by these physical and epidemiological facts. Instead, they are filtered through the complex lens of human perception, a process fraught with cognitive shortcuts (i.e. heuristics) that can systematically distort our understanding of risk.

These deviations from neoclassical economics are instances of what is known as behavioural economics. Neoclassical economics assumes rational actors with perfect information maximize personal utility, while behavioural economics incorporates psychology and accounts for real-world inconsistencies in rationality caused by cognitive biases, emotions, and context. Behavioral economics explains that people are often unable to fully analyse situations involving economic and probabilistic judgements. Individuals thus tend to rely on heuristics to form judgements, which can be systematically biased. (Kahneman and Smith, 2002)

This paper argues that to effectively address the air pollution crisis, we must move beyond neoclassical economic analysis and engage with the behavioral economics of risk perception, where heuristics play a decisive and often counterproductive role.

The imperative for this behavioral approach stems from the fundamental nature of the threat itself. The most harmful pollutants, such as fine particulate matter (PM_{2.5}), are imperceptible to the human sense. Unable to directly “see” the danger, individuals are forced to construct judgements based on imperfect proxies—visual haze, olfactory cues, media reports, or physical discomfort. This cognitive task, laden with uncertainty and complexity, triggers our reliance on intuitive mental shortcuts. These heuristics are essential for efficient decision-making but introduce predictable biases. In the context of an invisible, diffuse, and scientifically complex hazard like air pollution, the gap between objective data and subjective perception widens, making public comprehension and prioritization of the issue vulnerable to systematic error.

Understanding these cognitive processes is not a simple academic exercise but rather a critical component of policy design. If public concern, personal protective behavior, and political will are shaped more by heuristic-driven perceptions than by instrumental data, then policies based solely on perfect information or rational economic incentives may fail to resonate or motivate action. Behavioral economics provides the framework to decode this disconnect. By examining how specific biases operate within the domain of environmental risk, we can better anticipate public response, design more impactful communication strategies, and craft interventions that align with how people actually process information and make choices.

This paper proceeds by first establishing the necessary background on air pollution's physical characteristics and its documented health and economic impacts, while highlighting the inherent challenge of perceiving an invisible risk. It then delves into an analysis of three key heuristics—the halo effect, confirmation bias, and loss aversion—and their relevance to air pollution.

The background section outlines the objective reality of air pollution, citing recent global and national data that call for urgency, and detailing its detrimental health consequences. It also establishes why the perception of air pollution is a psychological puzzle, necessitating a behavioral inquiry.

The analysis on the halo effect explores how an overall positive (or negative) impression of a place, such as one's home or neighborhood, creates a cognitive bias that leads individuals to systematically overestimate (or underestimate) local air quality, even when faced with contrary objective data. Drawing on foundational behavioral economics studies and contemporary environmental research, such as investigations into wood-burning stove usage in Chile (Boso et al., 2020), this section demonstrates that overall neighbourhood satisfaction creates a “home halo”, leading residents to perceive indoor and local air as cleaner or dirtier than measurements indicate.

An examination of the confirmation bias investigates how individuals selectively interpret information about air pollution to reinforce pre-existing beliefs, leading to a persistent gap between improving objective data and stagnant or polarized public perception. The section dissects past studies on U.S. air quality trends and neighbourhood stigmas in Chicago to illustrate how deep-seated beliefs and social stigmas cause people to overlook empirical evidence of improvement or risk (King, 2014).

The discussion of loss aversion applies the principle of the disproportionate weight given to potential losses—compared to equivalent gains—to the slow adoption of pro-

environmental behavior and technologies. This section draws on experimental evidence and applied models, such as those analyzing wildfire risk perceptions and residential heating choices (Bartczak et al., 2015), to demonstrate that individuals' risk preferences become markedly more conservative when decisions involve perceptive losses, such as high upfront costs or lifestyle changes.

The paper proceeds as follows. In Section I, we delve deeper into the facts behind air pollution. In Sections II, III, and IV, we explore the halo effect, confirmation bias, and loss aversion respectively. Section V concludes.

Background

As touched upon in the introduction, air pollution represents one of the most significant environmental health risks of the modern era, with profound implications for public health and economic productivity. Simultaneously, perception of this risk is a complex psychological process, and may be filtered through cognitive biases before becoming a personalised human interpretation of objective reality. This section outlines the fundamental nature of air pollution as a physical phenomenon, its proven impacts, and the nuance involved in human perception, thereby establishing the necessity of investigating the heuristics that shape this perception.

Air pollution is the “contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere” (WHO, 2024). Key pollutants consist of particulate matter ($PM_{2.5}$ and PM_{10}), which can penetrate deep into the lungs and bloodstream; nitrogen dioxide (NO_2) and sulfur dioxide (SO_2), primarily from combustion processes; and ozone (O_3), a secondary pollutant formed in the atmosphere. Studies cited later on in this paper will focus on levels of $PM_{2.5}$ as a measure of air quality due to the seriousness of its health consequences and geographical ubiquity (Henning, 2024).

The World Health Organisation (WHO) has declared air pollution a public health emergency, with millions of premature deaths annually attributed to exposure (WHO, 2016). As a whole, air pollution accounts for 1 in 8 deaths around the world, the culprit of 7.9 million deaths in 2023 alone. More than 90% of East and Southeast Asia's population currently breathe air with unsafe levels of $PM_{2.5}$ and ground-level ozone (UNEP, 2024). According to the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) 2023, particulate matter pollution on its own was the second leading risk factor for early death and poor health globally (IHME, 2023). In particular, studies have shown $PM_{2.5}$ to increase risk of several diseases such as cardiovascular and respiratory diseases, lung cancer,

stroke, pneumonia, depression, and diabetes from short- and long-term exposure (Sukuman et al., 2023).

Apart from destructive health effects, air pollution also weighs heavily on countries' balance sheets. Estimates project a 2.9 trillion USD global economic cost, equivalent to 3.3% of the world's GDP. PM_{2.5} was found single-handedly responsible for 1.8 billion days of work absence, 4 million new cases of child asthma and 2 million preterm births in 2018 (Myllyvirta, 2020). In the UK, around 30,000 deaths per year are linked to air pollution, along with the depletion of £27 billion due to healthcare costs, productivity losses, and reduced quality of life, or closer to £50 billion if trickle-down effects like dementia are accounted for (RCP, 2025). The UK stands to benefit by £1.6 billion annually if WHO guidelines for air pollution are met (CBI, 2020). In India, air pollution cost an estimated \$95 billion (3% of the country's GDP) in the form of reduced productivity, work absences (3.8 billion working days), and premature deaths (responsible for 18% of deaths in India) in 2019. Had the country managed to half pollution in the last quarter of a decade, GDP is believed to will have risen by 4.5% (Dalberg, 2021). In Southeast Asia, deaths attributed to air pollution are projected to rise 10% by 2050, with costs snowballing to a hefty 600 billion USD. This figure captures indirect costs such as lost productivity, economic output, and premature deaths. Undoubtedly, good air quality is a key pillar of a strong economy.

While the physical and sociological facts of air pollution are critical, they do not, in themselves, dictate public response. The perception of air quality is inherently challenging because of the most harmful pollutants, such as PM_{2.5} which was termed in reference to its size of 2.5µm or smaller, are often invisible to the naked eye. Moreover, pollution-related health consequences often trigger delayed physical symptoms. In the absence of perfect sensory information, individuals cannot directly “see” the threat and must instead construct a judgement of an invisible environmental factor—air quality. (Fang et al., 2025)

To navigate a world full of such complex and uncertain judgements, the human brain relies on a powerful set of mental shortcuts known as heuristics or cognitive biases. Kahneman (2011) uses a conceptual framework of two “systems”—System 1 and System 2—in the human brain to explain this.¹⁵ This operation is a demonstration of our intuitive “System 1” in action, one half of two distinct cognitive systems involved in decision-making as introduced by Kahneman. The system is fast, automatic, and effortless, allowing us to make countless decisions without the laborious mental effort of our analytical “System 2”. Heuristics are the rules of thumb System 1 generally employs—they are generally effective but come with a trade-off: they can systematically lead us away from purely rational assessment and towards predictable errors.

These cognitive processes are not a sign of intellectual failure but rather a feature of our mental machinery evolved for efficiency. When faced with the daunting task of assessing an invisible danger like air pollution, it is cognitively efficient to substitute the hard question “What is the current concentration of PM_{2.5}?” with easier ones like “Does the air look hazy?” or “Does it smell bad?”. Common proxies for pollution levels include visual cues (haze or smog), olfactory cues (smell of smoke, exhaust, or industrial emissions), and contextual and somatic cues (media reports, seeing people wear masks, or experiencing physical symptoms like coughing or eye irritation). However, these proxies are imperfect and can be misleading—a clear, sunny day may still have dangerously high levels of ozone, while a foggy day might be mistaken for severe smog. This reliance on physical cues makes the perception of air quality highly susceptible to an unconscious use of cognitive biases as it facilitates a shift from a data-rich to cue-driven decision-making process and introduces perceptual vulnerabilities.

Consequently, the public’s perception of air quality is not a direct reflection of objective data but rather a heuristic-driven judgement, which in turn shapes personal behavior, public concern, and political will. In the following three sections, I further probe into three specific cognitive biases: the halo effect, confirmation bias, and loss aversion.

Halo Effect

If you were to meet someone for the first time and they treat you warmly and friendly, you would likely associate them with several other positive qualities, without any real evidence to support those assumptions. This is due to the human tendency to base our general judgement of a person or object off a single lasting impression of them. We can apply this in a general context where observation of a single positive or negative trait impels judgement in every other aspect. Formally, this is known as the ‘halo effect’.

The halo effect was first coined by American psychologist Eduard Thorndike in his personality studies and has since been one of the most widely known and discussed cognitive biases. Nisbett and Wilson (1977, Pg 1) outline the phenomenon as ‘the influence of a global evaluation on evaluations of individual attributes of a person (or object)’.¹⁶ We can understand the heuristic as a mental shortcut that causes an alteration in judgement when a positive or negative impression of a person, place, or object in one area influences our overall perception and assessment of them, even in unrelated areas.

Its existence has been tested and studied at length. Notable evidence comes from Nisbett and Wilson’s experiment (1977) on college students, to whom they played two different videotaped interviews of the same instructor. The experiment consisted of two

treatments—one where the instructor was warm and friendly, one where the instructor was cold and distant. After viewing, the students were asked to rate the instructor's likeability, as well as physical attractiveness, mannerisms and accent. The halo effect was measured quantitatively by calculating the difference in average ratings for each category between warm and cold treatments.

The between-subject study yielded more favourable ratings in all categories for the warm treatment and consistently poorer ratings for the cold treatment as displayed in the bar chart below.

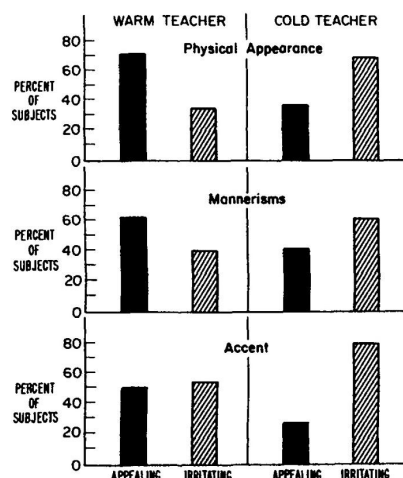


Figure 1: Ratings of Teacher's Physical Appearance, Mannerisms, and Accent as a Function of Manipulated Warmth vs Cold (Nisbett et al., 1977, Pg 253)

As the top third of Figure 1 shows, a sizable majority of warm treatment subjects rated the instructor's physical appearance as appealing, while a sizable majority of the cold treatment subjects rate the instructor's physical appearance as irritating. The central part of Figure 1 shows that ratings of the instructor's mannerisms are similar. Lastly, the bottom third of Figure 1 shows that about half of warm treatment subjects rated the instructor's accent as appealing, while an overwhelming majority of cold treatment subjects rated the instructor's accent as irritating. The results quantitatively prove the halo effect by showing a measurable shift in attribute judgment driven by an unconscious, global evaluation, even when objective information should allow for independent assessment.

Fundamentally, the halo effect causes our general impression of an entity to shape our view of its unknown attributes. This is especially relevant in air quality perception due to the ambiguity of air quality as an attribute. If one is emotionally attached to the place where they live (quantified by the difference in perception of air quality between home and city), the positive feelings evoked when thinking about their homes would subsequently

colour perceptions of air quality as clean and favourable (Hofflinger et al., 2019). Since global evaluations are powerful enough to trump other objective data available, the phenomenon likely explains why people in highly polluted cities still perceive their neighbourhood air quality positively. As a result, vigorous public awareness campaigns and a wide array of comprehensive research into the negative medical repercussions of poor air quality have yet to make a dent in prevailing high mortality rates due to air pollution. This highlights firsthand perception of surrounding air quality as the root of the issue. If people do not perceive the air they inhale as hazardous, they will not make an effort to rectify their habits to protect themselves (D' Antoni et al., 2017).

Previous research has examined the effect of the halo effect on individual perception of air pollution in specific contexts.

One such observational study was conducted in Southern Chile (Boso et al., 2020). As cited; wood-burning stoves are one of the main culprits of the high rates of pollution in countries across the Global South. Despite the presence of a wide range of technological options to replace stoves in Chile, it still exhibits similarly high levels of pollution. Outdoor and indoor pollution persists despite aggressive government intervention including measures such as refurbishment subsidies and prohibitions of firewood usage, suggesting the involvement of a halo effect. The study zeroes in on indoor air pollution to touch on reports that people tended to spend more than 80% of their time indoors.

Residents across 81 homes in Temuco and Padre Las Casas filled in questionnaires on their individual perception of air quality at home, which was then compared to on-site sensors and evaluations made by inspectors. The questionnaires consisted of 3 sections. First, a short five-item evaluating indoor and outdoor (neighbourhood and city) at that moment on a nine-point Likert scale. Then, an extended 25-item on air quality in general and on that day. Additional questions on severity of air pollution (five-point Likert scale), their level of concern for its effects on their health (five-point Likert scale), sociodemographic questions and home characteristics finished up the questionnaire.

The general environment of the study has high levels of particulate matter levels (PM level hereafter)—the recommended WHO level is $25\mu\text{g}/\text{m}^3$ while the detected indoor and outdoor $\text{PM}_{2.5}$ levels were $144\mu\text{g}/\text{m}^3$ and $74\mu\text{g}/\text{m}^3$ respectively. However, on a scale of 1-9, average evaluation of indoor air quality was relatively higher (3.97) than that of outdoor air quality (3.25). Residents also tended to evaluate indoor air quality as healthier than what the data showed, and more positive relative to inspectors. A 'Home Halo Effect' is measured by quantifying the difference between overestimation of indoor air quality in polluted households between residents and inspectors. Residents overestimate air quality by 80%

while inspectors only by 70%, meaning a 10% halo effect according to the authors' definition.

Building on the questionnaire data, Hofflinger (2019) sought to further establish the mechanisms that determine public perceptions of air quality as well as prove the prevalence of the halo effect over other related elements. Several independent factors—household income, education level, age, presence of respiratory ailments, number of wood-burning stoves, and whether they live in areas where wood-burning stove usage is restricted—were weighed against data collected on air quality evaluation in a mathematical model to investigate four dependent variables: perception of air quality in the home, neighbourhood, and the city, and the home halo effect. Results displayed a clear pattern of better subjective evaluation of air quality at home, as compared to the city or the neighbourhood—61% of respondents rated home air quality better than city, while only 7% thought it was worse. The mean value of the halo effect variable was positive, confirming a general trend of perceiving home air as cleaner than it is. These results are consistent with the “neighbourhood halo effect” (Bickerstaff and Walker, 2001) termed in accordance with evidence found that individuals generally perceived communities in close proximity to theirs to be less polluted than others. (Rankin, 1969)

Since the significance of the halo effect in public perception of air quality has been steadily supported, applying these insights to urban policy can potentially bring about improvements. After all, the public's understanding of and response to air pollution can have significant implications for the effectiveness of government interventions aimed at fostering the health and wellbeing of urban residents (Hofflinger et al., 2019). We could even go as far to say that the success of mitigation policies hinge on thorough understanding of people's attitudes towards indoor and outdoor air quality.

I suggest that governments can provide subsidies or tax rebates to real estate projects in fashionable neighbourhoods that meet a high threshold of integrated air-purifying green spaces into their buildings. By doing that, the government can incentivise a rise in green infrastructure in trendy urban areas. This would link pollution-reducing infrastructure with places that are seen as attractive, modern, or affluent, leveraging the halo effect to make sustainable features desirable status symbols. Residents would then begin to associate environmental improvement with aesthetic value, influencing them to incorporate it into their own home designs.

At the community level, local environmental agencies could partner with platforms to offer verified “Green-Status” badges on their user profiles, earned by documenting air-friendly actions (e.g. using public transport, installing an air-quality monitor). This allows environmental sustainability to be associated with social credibility and be cast under its

positive halo. Since people love to gain social capital and follow social trends, being eco-conscious will come to be seen as the “in” thing, encouraging widespread adoption of green behaviours through social proof.

Confirmation Bias

While most of us may consider ourselves open to learning and accepting new concepts and points of view, what we may not realise is our tendency to cherry-pick and interpret information in a manner that adheres to our existing beliefs and rejecting or discounting information that contradicts our beliefs. This phenomenon is attributed to “confirmation bias”, which, Nickerson (1998, Pg 1) defines as “the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand”.

Following in the footsteps of philosopher Karl Popper (1959), who believed hypotheses should be falsified instead of confirmed, Wason (1960) sought to experimentally demonstrate the existence of a confirmation bias. Participants in his experiment were given three numbers “2-4-6” and told that the triple satisfies a predetermined rule. Each round, they were then asked to construct a triple they believed fulfilled the rule as well, and subsequently feedback to whether the rule was fulfilled was given. When ready, participants were to state their assumed rule. Participants generally behaved in the following way: they formed a hypothesis on the rule (e.g. “consecutive even numbers”) then tested their hypothesis by proposing triples that satisfied the rule (e.g. “4-6-8”, “6-8-10”, “20-22-24”). The feedback was always positive. However, when participants ultimately presented their hypotheses after several rounds of testing, they were told they were wrong. In actuality, the rule was simply “ascending numbers”. The methods participants employed to test their hypothesis bore affirmative evidence despite their assumptions being wrong, exhibiting their use of a confirmation bias to arrive at their eventual conclusions (Oswald and Grosjean, 2004).

Expanding on this, Lord, Lee and Lepper (1981) examine how confirmation biases go further to increase belief polarisation—rather than decrease or leave it unaffected—when mixed or inconclusive findings are assimilated by proponents of opposite viewpoints.²⁵ Their study explores views on capital punishment between two groups of subjects recruited such that they were evenly split between being pro-deterrence and anti-deterrence. The study consists of two stages. In the first stage, subjects were given a brief overview of a study’s results, then asked to rate their change in attitudes toward capital punishment. In the second stage, subjects were provided with a detailed presentation of the procedure, results, and critiques, and again asked to evaluate its impact on their attitudes. This was then repeated with the second study. The first stage’s results reported proponents shifting

their views 3.75% towards becoming more in favour of capital punishment, whereas opponents shifted 3.13% towards being more opposed. Notably, the divergence widened in the second stage, with the beliefs of proponents and opponents strengthening in opposite directions by 9.38% and 10.6% respectively, despite substantial evidence being provided for both supporting and opposing views. Such data strongly supports that mixed data will lead to increased polarization rather than to uncertainty and moderation.

The confirmation bias is highly relevant in the context of air pollution, lying in the fact that air pollution has been a recognised problem for millennia (Fowler et al., 2020). Entrenched beliefs about atmospheric pollution, often solidified as social stigmas, can persist even when their original scientific basis becomes outdated (King, 2014; Shwartz, 2004). This persistence helps explain the limited impact of modern information campaigns, which seems to prompt limited environmental action (Ramirez et al., 2019).

Previous research, albeit sparse, substantiates the correlation between the confirmation bias and air pollution. Schwartz (2004) took an observational approach by comparing objective air pollution rates in the United States against subjective evaluation by citizens.²⁷ He cited public polls which found that the majority of Americans believed air pollution has gotten worse in the past 10 years (66%) and would continue to get worse (67%). On a state level, 78% of Californians were reported to believe the state has hardly made progress in solving environmental problems and around 80% of New Yorkers rate air pollution as a “very serious” or “somewhat serious” problem, as do 77% of Texans. Ironically enough, actual ozone level data reflect that the US has made dramatic progress in reducing air pollution over the last few decades, and most American cities now enjoy relatively good air quality. The nation’s monitoring stations registered ozone in excess of federal standards more than 12 times per year in the early 1980s, while the end of 2002 saw a fall to 4 exceedances per year. The considerable improvement in air quality should be perceptible, yet there remains a large disconnect between perception and reality. It is likely the presence of confirmation bias prompts individuals to cling to historically negative viewpoints of American air quality and overlook objective improvements, impeding their ability to perceive environmental impacts generated by measures put in place to mitigate air pollution.

An observational study by King (2014) probes into cases where confirmation bias contributes to false perceptions of air pollution due to influence from preconceived notions tied to socioeconomic levels.⁴ The study is set in Chicago, where—in contrast to the national trend—non-Hispanic Blacks are slightly less likely than other groups to live near toxic-releasing facilities and Hispanics are more likely to be exposed—consistent with a national tendency for higher environmental risk near concentrations of immigrants. Key results are as follows: residents who rated their air quality as “Poor” lived in

neighbourhoods where, on average, 21.5% of families were in poverty, conversely, those who rated their air as “Excellent” lived in area with only 10.9% poverty; “Poor” raters also lived in areas that were 37.9% Black and 33.6% Hispanic on average, whereas “Excellent” raters lived in areas that were only 22.5% Black and 20.9% Hispanic; 46.9% of residents with less than a high school education rated their air poorly, while only 28.4% of those with a college degree did the same. Despite the stark perceptual difference, the general difference in objective pollution levels between people who reported “Excellent” air and “Poor” air is actually small. Thus, social cues (in this case, racial demographics) emerge as the key predictors of air quality, a finding consistent with a neighborhood stigma interpretation.

King suggests a “neighbourhood stigma” theory wherein “once a place acquires a negative reputation, respondents will tend to rely heavily on that evaluation when inferring neighborhood characteristics, particularly in the absence of domain-specific information or experience”. While place reputation appears dependent on racial/ethnic phenotypes in Chicago, other considerations may be more salient in other areas. This relates back to confirmation bias by demonstrating how people are inclined to make judgements about air quality that align with preconceived social beliefs.

Fully grasping the implications and applications of confirmation bias can enable policy makers to not only alleviate its negative effects but harness it to positively influence public behaviour. Here I propose two communication strategies regulators can adopt to leverage on confirmation bias in the battle against air pollution.

The first strategy uses confirmation bias to link environmental messaging to identity and values. To do this, policymakers can frame pollution issues in a manner that aligns with different groups’ values. This plays into their confirmation biases and could potentially make them more receptive to proposed environmental actions. For instance, a family health and economic efficiency-oriented narrative can be channelled towards conservatives while climate justice and intergenerational responsibility can be emphasized to progressives. For a conservative audience, messaging could highlight how reducing air pollution protects children's health and lowers long-term healthcare costs. For a progressive audience, there could be a greater focus on the disproportionate impact poor air quality has on marginalised communities.

The second one involves curating source networks and information channels that hold sway over different groups. This approach leverages confirmation bias by strategically delivering the value-aligned messages from the first strategy through sources and channels the target audience already trusts and seeks out. For a conservative audience, this means partnering with and amplifying messages through faith leaders, military veterans’ groups,

and centre-right media personalities. For a progressive audience, effective channels include social justice organisations, educational institutions, and progressive media outlets.

Loss Aversion

A coin is tossed and the following question is posed to you: “If it’s tails, you lose \$10. How much would you have to gain on winning in order for this gamble to be acceptable to you?” Was your answer around \$20? That was the answer most of Kahneman’s students gave him. From this demonstration, we can plainly see that people place twice as much value on losing money as opposed to gaining money—a textbook example of the loss aversion hypothesis. (Kahneman, 2011)

The loss aversion hypothesis was first introduced by Kahneman and Tversky (1979) as an essential feature of prospect theory, which suggests people make decisions based on perceived value relative to a reference point instead of a static utility. The hypothesis is that “the response to losses is more extreme than the response to gain. The displeasure associated with losing a sum of money is generally greater than the pleasure associated with winning the same amount” (Kahneman and Tversky, 1981, Pg 454). A consequence of loss aversion is that it can lead to inconsistencies in risk preferences, i.e. the degree to which an individual prefers a sure amount to a gamble for any arbitrary risk. Essentially, an individual is risk-averse if they prefer the sure amount equal to the expected value of the gamble to the gamble, and risk-seeking if they prefer the latter. (Menezes and Hanson, 1970)

Karle et al. (2013) demonstrated the mechanisms of the heuristic in an experimental setting using choices among different lotteries.³² They gave subjects two series of choices, with 6 choices each, where they had to choose between lotteries and sure payments. For Series A, subjects had to make 6 choices between a lottery with 50% chance of winning 1 Euro and 50% chance of winning nothing, and a sure payment of S. S was either 10, 20, 30, 40, 50, or 60 Eurocents. This choice purely elicits their risk preferences without involving loss aversion as it only comprises gains. A risk-neutral individual would be indifferent between the lottery (which has an expected value of 50 Eurocents) and a sure payment of 50 Eurocents. Subjects who choose the lottery over sure payments of 50 Eurocents and above are risk-seekers, while those who choose sure payments of 50 Eurocents and below over the lottery are risk-aversers. In the experiment, 76.7% of participants pick a sure payment that is below or equal to 50 cents (risk-aversers).

In series B, subjects had to make six choices between a lottery that gave a $\frac{1}{3}$ chance of winning €1 and a $\frac{2}{3}$ chance of losing R Eurocents, and a sure payment of zero. R was either

0, 10, 20, 30, 50, 70, or 100 Eurocents. This choice is where loss aversion comes into play. Going by the same logic in Series A, a risk-neutral individual would be indifferent between a lottery of expected value zero (when $R=33$) and the sure payment. An individual that, additionally, displays loss aversion would be willing to accept a payment of zero more often for any given R . In the results of the experiment, 16.8% of participants make a choice that is more risk averse compared to what they did in Series A and therefore display a change in behaviour due to loss aversion.

Pollution links closely to loss aversion in terms of environmental action by individuals or groups. Whether or not to undertake a pro-environment action is a choice individuals have to make. That decision making process likely involves weighing private benefit derived against private cost personally incurred. Previous research done on the topic has shown that when faced with two possible choices, one energy-efficient and the other not, people will only choose the energy-efficient item if the payback time for the energy saved is very short (Kollmuss and Agyeman, 2002). This is consistent with loss averse behaviour as they value the loss of their time more than the environmental benefit. Moreover, since partaking in pro-environmental behaviour often requires a lifestyle change and a loss of the familiarity and safety in one's usual routine or incurs implementation costs, loss-averse individuals are likely to opt for the non-environmental option.

Bartczak et al. (2015) explores loss aversion in the environmental domain via a within-subject field study examining choices over a risky environmental good, wildfires in Poland.⁵ The study adapts Tanaka et al.'s (2010) experimental lottery design where Polish individuals are presented with three series of lottery pairs and asked to choose one lottery for each pair.³⁴ Respondents are asked to make a choice between programs that, instead of financial rewards, presented environmental effects of programs to reduce the risk of forest wildfires. In Treatment 1 and 2, they are told that there are two potential programs to further reduce the risk of forest fires which carry different success rates, and results are described as the annual average reduction of burned forest area in Poland and presented as uncertain. Simply put, the first two treatments involve choosing between only gains. As detailed above in Karle et al.'s study, this is used to determine risk preference independent of loss aversion. In Treatment 3 (the core treatment), respondents are asked to choose between programmes which would end in either success or failure, where more forests would be destroyed than currently. As there is now an element of loss, subjects' loss aversion can be elicited. Results from Treatment 1 and 2 indicated 75% of respondents were risk-averse. However, in Treatment 3 which introduces losses, the majority (~85%) of respondents exhibited risk-seeking behaviour instead. The paper provides conclusive evidence for loss aversion in the environmental context.

More recently, Knobloch et al.'s (2019) paper displays loss aversion's widespread applicability to the global energy model, using residential heating technology uptake as a case study.³⁵ The sector was chosen as decisions are overwhelmingly made by consumers, but not predominantly influenced by other interfering cultural factors. Additionally, the sector accounts for the largest share of direct greenhouse gas emissions by households. It is an observational study that performs analyses at both a global level and the level of individual countries, simulating future heating technology diffusion across 59 world regions.

The study employs a model that simulates consumer decision-making and weighs benefits (convenience, environmental benefits, etc.) and costs (upfront costs, operating costs, etc.). The model is run twice, with default loss aversion parameters and without. The model is used to compare the efficacy of four decarbonization policy scenarios with and without the effect of consumer loss aversion. These policy scenarios involve different levels of residential carbon taxes and subsidy payments. In this case, a carbon tax acts as the "gain" element as it increases the size of relative gains from switching to renewables by reducing energy costs, while subsidies act as the "loss" element as they reduce the size of relative losses. The aim is to assess how loss aversion influences model projections of technological change and policy effectiveness, especially as current energy models often assume optimal consumer decisions, which might lead to overly optimistic expectations of climate policy efficacy. Results indicate that loss-averse decision-makers are less likely to adopt energy-efficient technologies, as higher upfront costs are often perceived as losses with a stronger impact than future energy savings (evaluated as gains). This asymmetric valuation results in a "conservative shift" where preferences for current technologies are strengthened, reducing uptake of newer technologies.

The model further shows that subsidies were relatively more effective than carbon taxes as they directly address the pain point for loss-averse consumers—high investment costs for renewables. On the other hand, projections show that a carbon tax of 200 €/tCO₂ while considering loss aversion and a carbon tax of 100 €/tCO₂ without considering loss aversion achieve a similar reduction in overall emission levels. Thus, taxes require more drastic implementation, which is likely to be received poorly by the public. In short, the model tells us that we need double the tax to see plausible results comparable to idealistic projections (without loss aversion).

These differences in outcomes suggest that consumer behavior, specifically loss aversion, has a profound quantitative impact on the feasibility and required intensity of climate policies.

Having proven how loss aversion pertains to environmental pollution, policymakers can tap on this heuristic to maximise the effectiveness of policies aimed at reducing pollution.

While loss aversion is often discussed in the context of how it can hinder the effectiveness of policies, there are now several policies and interventions that have sought to actively “make use of” this behavioral trait to achieve desired outcomes. Current and previous approaches often involve framing choices to emphasize potential losses, thereby leveraging individuals’ stronger aversion to losing something compared to gaining an equivalent amount.

A notable application is in the design of incentives for healthy behaviours, where individuals are offered “insurance” (a sum of money) but told they will lose it if they do not achieve a certain healthy behaviour (e.g. attending an appointment or making a healthy choice). The prospect of losing this already “allocated” amount can be motivating and has been shown to encourage high rates of health behaviour uptake (Meeker et al., 2015).

This strategy can be adapted for pollution reduction in various ways, such as setting household carbon emission thresholds. A “green bonus” can be granted to households at the start of each month (e.g. credit on their utility bill) but lose part or all of it at the end of the month if they exceed the given emission threshold. This creates a stronger motivation to adopt greener appliances and habits than a standard rebate system.

Alternatively, Commuters could begin each week or month with a “streak bonus” (e.g. 10% discount on fares) that diminishes if they fall short of their weekly quota. In other words, missing a day of commuting via public transport would constitute a penalty (e.g. - 0.5% off the discount).

These examples serve to nudge individuals to more willingly adopt greener lifestyle habits.

Conclusion

This review has provided compelling evidence that heuristics—namely the halo effect, confirmation bias, and loss aversion—profoundly shape individual and collective perceptions of air quality, often creating a significant gap between objective reality and subjective interpretation. Findings explored in this paper are not merely curiosities of psychology but demonstrations of powerful, systematic forces that undermine the effectiveness of conventional pollution mitigation policies which assume rational, information-driven public behaviour.

The halo effect explains how positive emotional attachments to one’s living environment as a whole colours objective assessment of individual traits like air quality, leading to consistent overestimations of air quality that could create a false sense of security and

insulate people from the urgency of environmental risks. Confirmation bias leads communities to interpret ambiguous environmental cues through the lens of pre-existing beliefs and social stigmas, causing real improvements in air quality to have less than proportionate social impacts, and may even lead to ill-targeted policy actions. Loss aversion reveals a fundamental asymmetry in decision-making: the perceived pain of sacrificing convenience, money, or routine to adopt green behaviours far outweighs the perceived gain of future environmental benefits, creating a powerful inertia against pro-environmental action.

Crucially, these heuristics not only present obstacles but also offer novel pathways for policy innovation. Effective countervailing policies work with, not against these ingrained psychological patterns. As proposed, this can involve the following. We can leverage the halo effect by associating clean air technologies with socially desirable status and community identity. We can harness confirmation bias by framing environmental messages to align with deeply held values of different demographic groups, making sustainable choices feel like affirmations of identity instead of contradicting beliefs. Lastly, we can draw on people's loss aversion by restructuring incentives around the avoidance of loss—such as bonds forfeited for non-compliance with green policies.

Bridging the gap between air quality data and public perception to beget real, effective action requires a behavioural turn in environmental policy. We must go beyond mere information dissemination to strategically reshaping the cognitive and emotional landscape in which decisions are made. By integrating insights from behavioural science into policy design, regulators can craft more resonant interventions that align public perception with planetary health, ultimately transforming heuristics from barriers into catalysts for a cleaner future.

Bibliography

- (1) World Health Organization: WHO. (2019, July 30). Air pollution. https://www.who.int/health-topics/air-pollution#tab=tab_2
- (2) Kahneman, D., Smith, V. (2002) Advanced Information on the Prize in Economic Sciences. <https://www.nobelprize.org/uploads/2018/06/advanced-economicsciences2002-1.pdf>.
- (3) Boso, À., Álvarez, B., Oltra, C., Garrido, J., Muñoz, C., & Galvez-García, G. (2020). The grass is always greener on my side: A field experiment examining the home halo effect. *Sustainability*, 12(16), 6335. <https://doi.org/10.3390/su12166335>.

- (4) King, K. E. (2014). Chicago Residents' Perceptions of Air Quality: Objective Pollution, the Built Environment, and Neighborhood Stigma Theory. *Population and Environment*, 37(1), 1–21. <https://doi.org/10.1007/s11111-014-0228-x>
- (5) Bartczak, A., Chilton, S., & Meyerhoff, J. (2015). Wildfires in Poland: The impact of risk preferences and loss aversion on environmental choices. *Ecological Economics*, 116, 300–309. <https://doi.org/10.1016/j.ecolecon.2015.05.006>
- (6) Henning, R. J. (2024). Particulate Matter Air Pollution is a Significant Risk Factor for Cardiovascular Disease. *Current Problems in Cardiology*, 49(1, Part B), 102094. <https://doi.org/10.1016/j.cpcardiol.2023.102094>
- (7) UNEP. (2024). How Asian countries could save lives, boost growth by tackling air pollution. UNEP. <https://www.unep.org/news-and-stories/story/how-asian-countries-could-save-lives-boost-growth-tackling-air-pollution>
- (8) Institute for Health Metrics and Evaluation: IHME. (2023). Global Burden of Disease 2023: Findings from the GBD 2023 study. Institute for Health Metrics and Evaluation. <https://www.healthdata.org/research-analysis/library/global-burden-disease-2023-findings-gbd-2023-study>
- (9) Thanakon Sukuman, Ueda, K., Sarunya Sujaritpong, Hirunwut Praekunatham, Kornwipa Punnasiri, Tuangsit Wimuktayon, & Trakarn Prapasongsa. (2023). Health Impacts from PM2.5 Exposure Using Environmental Epidemiology and Health Risk Assessment: A Review. *Applied Environmental Research*, 1–14. <https://doi.org/10.35762/aer.2023010>
- (10) Myllyvirta, L. (2020). Quantifying the Economic Costs of Air Pollution from Fossil Fuels Key messages. Retrieved February 15, 2026, from <https://energyandcleanair.org/wp/wp-content/uploads/2020/02/Cost-of-fossil-fuels-briefing.pdf>.
- (11) Royal College of Physicians: RCP. (2025). A breath of fresh air: responding to the health challenges of modern air pollution. Rcp.ac.uk. <https://www.rcp.ac.uk/policy-and-campaigns/policy-documents/a-breath-of-fresh-air-responding-to-the-health-challenges-of-modern-air-pollution/>

- (12) CBI Economics. (2020). *Breathing life into the UK economy: Quantifying the economic benefits of cleaner air*. (2020). https://s40026.pcdn.co/wp-content/uploads/Breathing_life_into_the_UK_economy-Report-Final.pdf.
- (13) Berg, L. (2021, September). More than a Public Health Crisis: Air Pollution's Impact on the Indian Economy. *Dalberg*. <https://dalberg.com/our-ideas/more-than-a-public-health-crisis-air-pollutions-impact-on-the-indian-economy/>
- (14) Fang, T., Gu, Y., Li, J., Sung, J. J. Y., Yim, S. H. L. (2025). Effects of Climate Change on PM2.5 Formation Chemistry, and the Associated Human Health Impacts and Economic Costs in Southeast Asia. *Environment International*, 205, 109835. <https://doi.org/10.1016/j.envint.2025.109835>
- (15) Kahneman, D. (2011). *Thinking, Fast and Slow*. Penguin.
- (16) Nisbett, R. E., & Wilson, T. D. (1977). The Halo Effect: Evidence for Unconscious Alteration of Judgments. *Journal of Personality and Social Psychology*, 35(4), 250–256. <https://doi.org/10.1037/0022-3514.35.4.250>
- (17) Hofflinger, Á., Boso, À., & Oltra, C. (2019). The Home Halo Effect: how Air Quality Perception is Influenced by Place Attachment. *Human Ecology*, 47(4), 589–600. <https://doi.org/10.1007/s10745-019-00100-z>
- (18) D'Antoni, D., Smith, L., Auyeung, V., & Weinman, J. (2017). Psychosocial and Demographic Predictors of Adherence and Non-adherence to Health Advice Accompanying Air Quality Warning Systems: A Systematic Review. *Environmental Health*, 16(1). <https://doi.org/10.1186/s12940-017-0307-4>
- (19) Bickerstaff, K., & Walker, G. (2001). Public understandings of air pollution: the “localisation” of environmental risk. *Global Environmental Change*, 11(2), 133–145. [https://doi.org/10.1016/S0959-3780\(00\)00063-7](https://doi.org/10.1016/S0959-3780(00)00063-7)
- (20) Rankin, R. E. (1969). Air Pollution Control and Public Apathy. *Journal of the Air Pollution Control Association*, 19(8), 565–569. <https://doi.org/10.1080/00022470.1969.10466523>

- (21) Nickerson, R. S. (1998). Confirmation bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- (22) Popper, K. R. (1977). The Logic of Scientific Discovery. *Systematic Zoology*, 26(3), 361. <https://doi.org/10.2307/2412687>
- (23) Wason, P. (1960). On the Failure to Eliminate Hypotheses in a Conceptual Task. *Quarterly Journal of Experimental Psychology*, 12(3), 129–140. <https://doi.org/10.1080/17470216008416717>
- (24) Oswald, M. E., Grosjean, S. (2012). Confirmation bias. In R. F. Pohl (Eds.), *Cognitive Illusions. A Handbook on Fallacies and Biases in Thinking, Judgement and Memory* (pp 91–108). Psychology Press.
- (25) Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098–2109.
- (26) Fowler, D. (2020). A Chronology of Global Air Quality. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 378(2183). <https://doi.org/10.1098/rsta.2019.0314>
- (27) Schwartz, J. (2024). *Clearing the Air*. Ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=505604
- (28) Ramírez, A. S., Ramondt, S., Van Bogart, K., & Perez-Zuniga, R. (2019). Public Awareness of Air Pollution and Health Threats: Challenges and Opportunities for Communication Strategies To Improve Environmental Health Literacy. *Journal of Health Communication*, 24(1), 75–83. <https://doi.org/10.1080/10810730.2019.1574320>
- (29) Kahneman, D., & Tversky, A. (1979). Prospect theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2), 263–292. <https://doi.org/10.2307/1914185>

- (30) Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 453–458.
<https://doi.org/10.1126/science.7455683>
- (31) Menezes, C. F., & Hanson, D. L. (1970). On the Theory of Risk Aversion. *International Economic Review*, 11(3), 481. <https://doi.org/10.2307/2525326>
- (32) Karle, H., Kirchsteiger, G., & Peitz, M. (2015). Loss Aversion and Consumption Choice: Theory and Experimental Evidence. *American Economic Journal: Microeconomics*, 7(2), 101–120. <https://doi.org/10.1257/mic.20130104>
- (33) Kollmuss, A., & Agyeman, J. (2002). Mind the Gap: Why Do People Act Environmentally and What Are the Barriers to pro-environmental behavior? *Environmental Education Research*, 8(3), 239–260.
<https://doi.org/10.1080/13504620220145401>
- (34) Tanaka, T., Camerer, C. F., & Nguyen, Q. (2006). Risk and Time Preferences: Experimental and Household Survey Data from Vietnam. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.877229>
- (35) Knobloch, F., Huijbregts, M. A. J., & Mercure, J.-F. (2019). Modelling the effectiveness of climate policies: How important is loss aversion by consumers? *Renewable and Sustainable Energy Reviews*, 116, 109419.
<https://doi.org/10.1016/j.rser.2019.109419>
- (36) Meeker, D., Thompson, C., Strylewicz, G., Knight, T. K., & Doctor, J. N. (2015). Use of Insurance Against a Small Loss as an Incentive Strategy. *Decision Analysis*, 12(3), 122–129. <https://doi.org/10.1287/deca.2015.0314>

The paper provides interesting and potentially valuable insights by compiling findings from existing studies based on previously published research. The topic is relevant and timely; however, in its current form, the manuscript lacks methodological and conceptual strength in how the findings are presented and justified. Substantial revisions are required to improve clarity, rigour, and scholarly contribution.

Major Comments

1. Methodology and Study Selection (Major Concern)

The most significant limitation of the paper is the absence of a clearly defined and transparent methodology. As the study is based on secondary data and published literature, it is essential to explicitly describe:

- The criteria used for selecting studies (e.g. inclusion and exclusion criteria)
- The sources/databases consulted
- The timeframe covered by the literature
- The data extraction process, including how variables or indicators were identified
- Any data frameworks, classification systems, or analytical approaches used to synthesise findings

Without this information, it is difficult to assess the credibility, comprehensiveness, and reproducibility of the review. The author(s) should clarify whether the approach follows a systematic, scoping, or narrative review framework and justify the chosen method.

2. The introduction should more explicitly highlight what is missing in current research, how this paper responds to that gap, and why this contribution is novel. While the paper draws on existing studies, the value added by this review is not sufficiently clear. Greater emphasis is needed on how this work offers new insights, frameworks, or interpretations.
3. The results and discussion sections require stronger integration and a more analytical presentation. Specifically, findings should be systematically organised, potentially through thematic synthesis or analytical categorisation

Stronger connections between the findings and the study objectives, as well as existing theory or policy contexts, would significantly enhance the paper.

Overall Recommendation:

The paper has potential and addresses an important topic; however, major revisions are required before it can be considered for publication. Strengthening the methodology, clarifying the study's rationale and novelty, and presenting results and discussion in a more analytical and innovative manner will substantially improve the quality and impact of the manuscript. The manuscript would benefit from improved clarity in language and structure in several sections.

The Psychology of Pollution: Analysing How Heuristics Affect Perception of Air Quality and Policy Design

[Student name redacted]

[School name redacted]

Keywords:

Pollution, Air quality, Behavioural economics, Confirmation bias, Halo effect, Loss aversion, Policy Design

Abstract

Air pollution is a severe, global issue responsible for millions of premature deaths yearly and economic costs totaling trillions of dollars. Heuristics—mental shortcuts individuals intuitively employ to aid decision-making—are found to affect public perception of this issue due its subjectivity and complexity. This paper conducts an analysis into the critical role of three key heuristics (the halo effect, confirmation bias, and loss aversion) in shaping public perception, often leading to a significant gap between objective reality and subjective experience. The halo effect causes residents to systematically overestimate the cleanliness of air in their own homes due to emotional place attachment shaping global evaluations of their living environments. Confirmation bias explains the persistence of entrenched negative perceptions of air pollution in the face of objective environmental improvements. Loss aversion hinders the adoption of energy-efficient technologies and reduces the effectiveness of policies that trigger individuals' loss aversion, such as carbon taxes. Through a review of foundational psychological experiments and environmental studies, this analysis demonstrates how these heuristics systematically distort risk assessment and hinder effective public response to pollution. The paper concludes by proposing targeted urban, communication, and incentive policies designed to either mitigate and/or strategically harness these biases.

The Psychology of Pollution: Analysing How Heuristics Affect Perception of Air Quality and Policy Design

I. Introduction

Air pollution stands as one of the most pervasive and severe environmental causes of the 21st century, imposing a staggering toll on global public health and economic stability (WHO, 2016). Millions of premature deaths annually are attributed to exposure (WHO, 2016). Economic costs amounting to trillions of dollars—eroding national GDPs and straining healthcare systems—are starkly documented. Yet, societal and individual responses to this singular threat are not dictated solely by these physical and epidemiological facts. Instead, they are filtered through the complex lens of human perception, a process fraught with cognitive shortcuts (i.e. heuristics) that can systematically distort our understanding of risk.

These deviations from neoclassical economics are instances of what is known as behavioural economics. Neoclassical economics assumes rational actors with perfect information maximize personal utility, while behavioural economics incorporates psychology and accounts for real-world inconsistencies in rationality caused by cognitive biases, emotions, and context. Behavioral economics explains that people are often unable to fully analyse situations involving economic and probabilistic judgements. Individuals thus tend to rely on heuristics to form judgements, which can be systematically biased. (Kahneman and Smith, 2002)

This paper argues that to effectively address the air pollution crisis, we must move beyond neoclassical economic analysis and engage with the behavioral economics of risk perception, where heuristics play a decisive and often counterproductive role.

The imperative for this behavioral approach stems from the fundamental nature of the threat itself. The most harmful pollutants, such as fine particulate matter (PM_{2.5}), are imperceptible to the human sense. Unable to directly “see” the danger, individuals are forced to construct judgements based on imperfect proxies—visual haze, olfactory cues, media reports, or physical discomfort. This cognitive task, laden with uncertainty and complexity, triggers our reliance on intuitive mental shortcuts. These heuristics are essential for efficient decision-making but introduce predictable biases. In the context of an invisible, diffuse, and scientifically complex hazard like air pollution, the gap between objective data and subjective perception widens, making public comprehension and prioritization of the issue vulnerable to systematic error.

Understanding these cognitive processes is not a simple academic exercise but rather a critical component of policy design. If public concern, personal protective behavior, and political will are shaped more by heuristic-driven perceptions than by instrumental data, then policies based solely on perfect information or rational economic incentives may fail to resonate or motivate action. Behavioral economics provides the framework to decode this disconnect. By examining how specific biases operate within the domain of environmental risk, we can better anticipate public response, design more impactful communication strategies, and craft interventions that align with how people actually process information and make choices.

There exists a growing body of literature on both air pollution and behavioural economics individually, but a critical gap persists at their intersection. Where behavioural approaches have been recently applied to environmental risk perception, they tend to examine cognitive biases within a narrow empirical scope, rather than developing an integrated account from various contexts of how the general bias impacts air quality perception as a whole. The systemic application of heuristics—namely the halo effect confirmation bias, and loss aversion—to air quality perception and pollution-related policy design therefore remains underdeveloped in existing literature. This paper responds to that gap by examining each of these three heuristics in turn, synthesising psychological and environmental evidence to demonstrate how each independently distorts public understanding of air quality and undermines conventional policy responses. z

This paper proceeds by first establishing the necessary background on air pollution's physical characteristics and its documented health and economic impacts, while highlighting the inherent challenge of perceiving an invisible risk. It then delves into an analysis of three key heuristics—the halo effect, confirmation bias, and loss aversion—and their relevance to air pollution.

The background section outlines the objective reality of air pollution, citing recent global and national data that call for urgency, and detailing its detrimental health consequences. It also establishes why the perception of air pollution is a psychological puzzle, necessitating a behavioral inquiry.

The analysis on the halo effect explores how an overall positive (or negative) impression of a place, such as one's home or neighborhood, creates a cognitive bias that leads individuals to systematically overestimate (or underestimate) local air quality, even when faced with contrary objective data. Drawing on foundational behavioral economics studies and contemporary environmental research, such as investigations into wood-burning stove usage in Chile (Boso et al., 2020), this section demonstrates that overall neighbourhood

satisfaction creates a “home halo”, leading residents to perceive indoor and local air as cleaner or dirtier than measurements indicate.

An examination of the confirmation bias investigates how individuals selectively interpret information about air pollution to reinforce pre-existing beliefs, leading to a persistent gap between improving objective data and stagnant or polarized public perception. The section dissects past studies on U.S. air quality trends and neighbourhood stigmas in Chicago to illustrate how deep-seated beliefs and social stigmas cause people to overlook empirical evidence of improvement or risk (King, 2014).

The discussion of loss aversion applies the principle of the disproportionate weight given to potential losses—compared to equivalent gains—to the slow adoption of pro-environmental behavior and technologies. This section draws on experimental evidence and applied models, such as those analyzing wildfire risk perceptions and residential heating choices (Bartczak et al., 2015), to demonstrate that individuals’ risk preferences become markedly more conservative when decisions involve perceptive losses, such as high upfront costs or lifestyle changes.

The paper proceeds as follows. In Section I, we clarify the methodology for study selection and data extraction. In Section II, we delve deeper into the facts behind air pollution. In Sections II, III, and IV, we explore the halo effect, confirmation bias, and loss aversion respectively. Section V concludes.

II. Methodology

This paper adopts a narrative review framework to synthesise existing empirical and theoretical literature on the role of cognitive heuristics in shaping public perception of air quality. A narrative approach was selected on the grounds that the primary objective of this review is thematic synthesis and theoretical integration across disciplines, namely behavioural economics, environmental science, and public policy, rather than an exhaustive enumeration of all available studies.

Literature was identified through searches of Google Scholar, PubMed, Web of Science, and JSTOR, supplemented by reports from international organisations including the World Organisation (WHO), the United Nations Environment Programme (UNEP), and the Institute for Health Metrics and Evaluation (IHME). Search terms included combinations of the following: “air pollution”, “halo effect”, “confirmation bias”, “loss aversion”, “air quality perception”, and “cognitive bias”. No strict date restriction was imposed, as the review deliberately includes foundational experimental studies and theoretical papers, such as

Kahneman and Tversky's on prospect theory (1979) and Wason's on the rule discovery task (1960), providing a core, well-tested psychological base to build upon. More recent empirical studies, prioritised from 2000 onwards, were selected to ground these psychological frameworks in contemporary environmental contexts.

Studies were included based on relevance to at least one of the three heuristics under examination, methodological transparency and publication in peer-reviewed journals or by credible international institutions; purely anecdotal or non-empirical commentary was excluded. Findings were then organised thematically; for each heuristic, foundational experimental evidence establishing the psychological mechanism was identified first, followed by domain-specific studies applying that mechanism to air pollution perception or responsive behaviour. Policy implications were derived inductively from the patterns observed across studies within each thematic cluster, with the aim of translating behavioural insights into actionable regulatory recommendations.

III. Background

As touched upon in the introduction, air pollution represents one of the most significant environmental health risks of the modern era, with profound implications for public health and economic productivity. Simultaneously, perception of this risk is a complex psychological process, and may be filtered through cognitive biases before becoming a personalised human interpretation of objective reality. This section outlines the fundamental nature of air pollution as a physical phenomenon, its proven impacts, and the nuance involved in human perception, thereby establishing the necessity of investigating the heuristics that shape this perception.

Air pollution is the “contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere” (WHO, 2024). Key pollutants consist of particulate matter ($PM_{2.5}$ and PM_{10}), which can penetrate deep into the lungs and bloodstream; nitrogen dioxide (NO_2) and sulfur dioxide (SO_2), primarily from combustion processes; and ozone (O_3), a secondary pollutant formed in the atmosphere. Studies cited later on in this paper will focus on levels of $PM_{2.5}$ as a measure of air quality due to the seriousness of its health consequences and geographical ubiquity (Henning, 2024).

The World Health Organisation (WHO) has declared air pollution a public health emergency, with millions of premature deaths annually attributed to exposure (WHO, 2016). As a whole, air pollution accounts for 1 in 8 deaths around the world, the culprit of 7.9 million deaths in 2023 alone. More than 90% of East and Southeast Asia's population

currently breathe air with unsafe levels of PM_{2.5} and ground-level ozone (UNEP, 2024). According to the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) 2023, particulate matter pollution on its own was the second leading risk factor for early death and poor health globally (IHME, 2023). In particular, studies have shown PM_{2.5} to increase risk of several diseases such as cardiovascular and respiratory diseases, lung cancer, stroke, pneumonia, depression, and diabetes from short- and long-term exposure (Sukuman et al., 2023).

Apart from destructive health effects, air pollution also weighs heavily on countries' balance sheets. Estimates project a 2.9 trillion USD global economic cost, equivalent to 3.3% of the world's GDP. PM_{2.5} was found single-handedly responsible for 1.8 billion days of work absence, 4 million new cases of child asthma and 2 million preterm births in 2018 (Myllyvirta, 2020). In the UK, around 30,000 deaths per year are linked to air pollution, along with the depletion of £27 billion due to healthcare costs, productivity losses, and reduced quality of life, or closer to £50 billion if trickle-down effects like dementia are accounted for (RCP, 2025). The UK stands to benefit by £1.6 billion annually if WHO guidelines for air pollution are met (CBI, 2020). In India, air pollution cost an estimated \$95 billion (3% of the country's GDP) in the form of reduced productivity, work absences (3.8 billion working days), and premature deaths (responsible for 18% of deaths in India) in 2019. Had the country managed to half pollution in the last quarter of a decade, GDP is believed to will have risen by 4.5% (Dalberg, 2021). In Southeast Asia, deaths attributed to air pollution are projected to rise 10% by 2050, with costs snowballing to a hefty 600 billion USD. This figure captures indirect costs such as lost productivity, economic output, and premature deaths. Undoubtedly, good air quality is a key pillar of a strong economy.

While the physical and sociological facts of air pollution are critical, they do not, in themselves, dictate public response. The perception of air quality is inherently challenging because of the most harmful pollutants, such as PM_{2.5} which was termed in reference to its size of 2.5µm or smaller, are often invisible to the naked eye. Moreover, pollution-related health consequences often trigger delayed physical symptoms. In the absence of perfect sensory information, individuals cannot directly “see” the threat and must instead construct a judgement of an invisible environmental factor—air quality. (Fang et al., 2025)

To navigate a world full of such complex and uncertain judgements, the human brain relies on a powerful set of mental shortcuts known as heuristics or cognitive biases. Kahneman (2011) uses a conceptual framework of two “systems”—System 1 and System 2—in the human brain to explain this.¹⁵ This operation is a demonstration of our intuitive “System 1” in action, one half of two distinct cognitive systems involved in decision-making as introduced by Kahneman. The system is fast, automatic, and effortless, allowing us to make countless decisions without the laborious mental effort of our analytical “System 2”.

Heuristics are the rules of thumb System 1 generally employs—they are generally effective but come with a trade-off: they can systematically lead us away from purely rational assessment and towards predictable errors.

These cognitive processes are not a sign of intellectual failure but rather a feature of our mental machinery evolved for efficiency. When faced with the daunting task of assessing an invisible danger like air pollution, it is cognitively efficient to substitute the hard question “What is the current concentration of PM_{2.5}?” with easier ones like “Does the air look hazy?” or “Does it smell bad?”. Common proxies for pollution levels include visual cues (haze or smog), olfactory cues (smell of smoke, exhaust, or industrial emissions), and contextual and somatic cues (media reports, seeing people wear masks, or experiencing physical symptoms like coughing or eye irritation). However, these proxies are imperfect and can be misleading—a clear, sunny day may still have dangerously high levels of ozone, while a foggy day might be mistaken for severe smog. This reliance on physical cues makes the perception of air quality highly susceptible to an unconscious use of cognitive biases as it facilitates a shift from a data-rich to cue-driven decision-making process and introduces perceptual vulnerabilities.

Consequently, the public’s perception of air quality is not a direct reflection of objective data but rather a heuristic-driven judgement, which in turn shapes personal behavior, public concern, and political will. In the following three sections, I further probe into three specific cognitive biases: the halo effect, confirmation bias, and loss aversion.

IV. Halo Effect

If you were to meet someone for the first time and they treat you warmly and friendly, you would likely associate them with several other positive qualities, without any real evidence to support those assumptions. This is due to the human tendency to base our general judgement of a person or object off a single lasting impression of them. We can apply this in a general context where observation of a single positive or negative trait impels judgement in every other aspect. Formally, this is known as the ‘halo effect’.

The halo effect was first coined by American psychologist Eduard Thorndike in his personality studies and has since been one of the most widely known and discussed cognitive biases. Nisbett and Wilson (1977, Pg 1) outline the phenomenon as ‘the influence of a global evaluation on evaluations of individual attributes of a person (or object)’.¹⁶ We can understand the heuristic as a mental shortcut that causes an alteration in judgement when a positive or negative impression of a person, place, or object in one area influences our overall perception and assessment of them, even in unrelated areas.

Its existence has been tested and studied at length. Notable evidence comes from Nisbett and Wilson's experiment (1977) on college students, to whom they played two different videotaped interviews of the same instructor. The experiment consisted of two treatments—one where the instructor was warm and friendly, one where the instructor was cold and distant. After viewing, the students were asked to rate the instructor's likeability, as well as physical attractiveness, mannerisms and accent. The halo effect was measured quantitatively by calculating the difference in average ratings for each category between warm and cold treatments.

The between-subject study yielded more favourable ratings in all categories for the warm treatment and consistently poorer ratings for the cold treatment as displayed in the bar chart below.

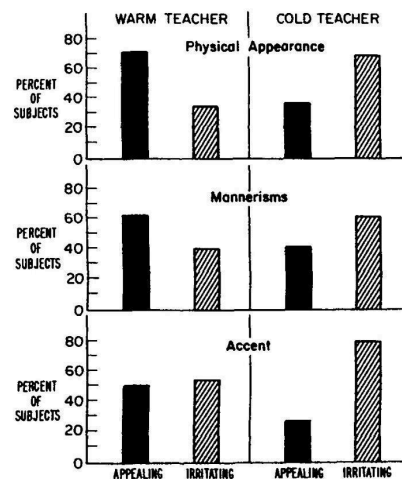


Figure 1: Ratings of Teacher's Physical Appearance, Mannerisms, and Accent as a Function of Manipulated Warmth vs Cold (Nisbett et al., 1977, Pg 253)

As the top third of Figure 1 shows, a sizable majority of warm treatment subjects rated the instructor's physical appearance as appealing, while a sizable majority of the cold treatment subjects rate the instructor's physical appearance as irritating. The central part of Figure 1 shows that ratings of the instructor's mannerisms are similar. Lastly, the bottom third of Figure 1 shows that about half of warm treatment subjects rated the instructor's accent as appealing, while an overwhelming majority of cold treatment subjects rated the instructor's accent as irritating. The results quantitatively prove the halo effect by showing a measurable shift in attribute judgment driven by an unconscious, global evaluation, even when objective information should allow for independent assessment.

Fundamentally, the halo effect causes our general impression of an entity to shape our view of its unknown attributes. This is especially relevant in air quality perception due to the ambiguity of air quality as an attribute. If one is emotionally attached to the place where they live (quantified by the difference in perception of air quality between home and city), the positive feelings evoked when thinking about their homes would subsequently colour perceptions of air quality as clean and favourable (Hofflinger et al., 2019). Since global evaluations are powerful enough to trump other objective data available, the phenomenon likely explains why people in highly polluted cities still perceive their neighbourhood air quality positively. As a result, vigorous public awareness campaigns and a wide array of comprehensive research into the negative medical repercussions of poor air quality have yet to make a dent in prevailing high mortality rates due to air pollution. This highlights firsthand perception of surrounding air quality as the root of the issue. If people do not perceive the air they inhale as hazardous, they will not make an effort to rectify their habits to protect themselves (D' Antoni et al., 2017).

Previous research has examined the effect of the halo effect on individual perception of air pollution in specific contexts.

One such observational study was conducted in Southern Chile (Boso et al., 2020). As cited; wood-burning stoves are one of the main culprits of the high rates of pollution in countries across the Global South. Despite the presence of a wide range of technological options to replace stoves in Chile, it still exhibits similarly high levels of pollution. Outdoor and indoor pollution persists despite aggressive government intervention including measures such as refurbishment subsidies and prohibitions of firewood usage, suggesting the involvement of a halo effect. The study zeroes in on indoor air pollution to touch on reports that people tended to spend more than 80% of their time indoors.

Residents across 81 homes in Temuco and Padre Las Casas filled in questionnaires on their individual perception of air quality at home, which was then compared to on-site sensors and evaluations made by inspectors. The questionnaires consisted of 3 sections. First, a short five-item evaluating indoor and outdoor (neighbourhood and city) at that moment on a nine-point Likert scale. Then, an extended 25-item on air quality in general and on that day. Additional questions on severity of air pollution (five-point Likert scale), their level of concern for its effects on their health (five-point Likert scale), sociodemographic questions and home characteristics finished up the questionnaire.

The general environment of the study has high levels of particulate matter levels (PM level hereafter)—the recommended WHO level is $25\mu\text{g}/\text{m}^3$ while the detected indoor and outdoor $\text{PM}_{2.5}$ levels were $144\mu\text{g}/\text{m}^3$ and $74\mu\text{g}/\text{m}^3$ respectively. However, on a scale of 1-9, average evaluation of indoor air quality was relatively higher (3.97) than that of outdoor air

quality (3.25). Residents also tended to evaluate indoor air quality as healthier than what the data showed, and more positive relative to inspectors. A 'Home Halo Effect' is measured by quantifying the difference between overestimation of indoor air quality in polluted households between residents and inspectors. Residents overestimate air quality by 80% while inspectors only by 70%, meaning a 10% halo effect according to the authors' definition.

Building on the questionnaire data, Hofflinger (2019) sought to further establish the mechanisms that determine public perceptions of air quality as well as prove the prevalence of the halo effect over other related elements. Several independent factors—household income, education level, age, presence of respiratory ailments, number of wood-burning stoves, and whether they live in areas where wood-burning stove usage is restricted—were weighed against data collected on air quality evaluation in a mathematical model to investigate four dependent variables: perception of air quality in the home, neighbourhood, and the city, and the home halo effect. Results displayed a clear pattern of better subjective evaluation of air quality at home, as compared to the city or the neighbourhood—61% of respondents rated home air quality better than city, while only 7% thought it was worse. The mean value of the halo effect variable was positive, confirming a general trend of perceiving home air as cleaner than it is. These results are consistent with the “neighbourhood halo effect” (Bickerstaff and Walker, 2001) termed in accordance with evidence found that individuals generally perceived communities in close proximity to theirs to be less polluted than others. (Rankin, 1969)

Since the significance of the halo effect in public perception of air quality has been steadily supported, applying these insights to urban policy can potentially bring about improvements. After all, the public's understanding of and response to air pollution can have significant implications for the effectiveness of government interventions aimed at fostering the health and wellbeing of urban residents (Hofflinger et al., 2019). We could even go as far to say that the success of mitigation policies hinge on thorough understanding of people's attitudes towards indoor and outdoor air quality.

I suggest that governments can provide subsidies or tax rebates to real estate projects in fashionable neighbourhoods that meet a high threshold of integrated air-purifying green spaces into their buildings. By doing that, the government can incentivise a rise in green infrastructure in trendy urban areas. This would link pollution-reducing infrastructure with places that are seen as attractive, modern, or affluent, leveraging the halo effect to make sustainable features desirable status symbols. Residents would then begin to associate environmental improvement with aesthetic value, influencing them to incorporate it into their own home designs.

At the community level, local environmental agencies could partner with platforms to offer verified “Green-Status” badges on their user profiles, earned by documenting air-friendly actions (e.g. using public transport, installing an air-quality monitor). This allows environmental sustainability to be associated with social credibility and be cast under its positive halo. Since people love to gain social capital and follow social trends, being eco-conscious will come to be seen as the “in” thing, encouraging widespread adoption of green behaviours through social proof.

V. Confirmation Bias

While most of us may consider ourselves open to learning and accepting new concepts and points of view, what we may not realise is our tendency to cherry-pick and interpret information in a manner that adheres to our existing beliefs and rejecting or discounting information that contradicts our beliefs. This phenomenon is attributed to “confirmation bias”, which, Nickerson (1998, Pg 1) defines as “the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand”.

Following in the footsteps of philosopher Karl Popper (1959), who believed hypotheses should be falsified instead of confirmed, Wason (1960) sought to experimentally demonstrate the existence of a confirmation bias. Participants in his experiment were given three numbers “2-4-6” and told that the triple satisfies a predetermined rule. Each round, they were then asked to construct a triple they believed fulfilled the rule as well, and subsequently feedback to whether the rule was fulfilled was given. When ready, participants were to state their assumed rule. Participants generally behaved in the following way: they formed a hypothesis on the rule (e.g. “consecutive even numbers”) then tested their hypothesis by proposing triples that satisfied the rule (e.g. “4-6-8”, “6-8-10”, “20-22-24”). The feedback was always positive. However, when participants ultimately presented their hypotheses after several rounds of testing, they were told they were wrong. In actuality, the rule was simply “ascending numbers”. The methods participants employed to test their hypothesis bore affirmative evidence despite their assumptions being wrong, exhibiting their use of a confirmation bias to arrive at their eventual conclusions (Oswald and Grosjean, 2004).

Expanding on this, Lord, Lee and Lepper (1981) examine how confirmation biases go further to increase belief polarisation—rather than decrease or leave it unaffected—when mixed or inconclusive findings are assimilated by proponents of opposite viewpoints.²⁵ Their study explores views on capital punishment between two groups of subjects recruited such that they were evenly split between being pro-deterrence and anti-deterrence. The study consists of two stages. In the first stage, subjects were given a brief overview of a study’s

results, then asked to rate their change in attitudes toward capital punishment. In the second stage, subjects were provided with a detailed presentation of the procedure, results, and critiques, and again asked to evaluate its impact on their attitudes. This was then repeated with the second study. The first stage's results reported proponents shifting their views 3.75% towards becoming more in favour of capital punishment, whereas opponents shifted 3.13% towards being more opposed. Notably, the divergence widened in the second stage, with the beliefs of proponents and opponents strengthening in opposite directions by 9.38% and 10.6% respectively, despite substantial evidence being provided for both supporting and opposing views. Such data strongly supports that mixed data will lead to increased polarization rather than to uncertainty and moderation.

The confirmation bias is highly relevant in the context of air pollution, lying in the fact that air pollution has been a recognised problem for millennia (Fowler et al., 2020). Entrenched beliefs about atmospheric pollution, often solidified as social stigmas, can persist even when their original scientific basis becomes outdated (King, 2014; Schwartz, 2004). This persistence helps explain the limited impact of modern information campaigns, which seems to prompt limited environmental action (Ramirez et al., 2019).

Previous research, albeit sparse, substantiates the correlation between the confirmation bias and air pollution. Schwartz (2004) took an observational approach by comparing objective air pollution rates in the United States against subjective evaluation by citizens.²⁷ He cited public polls which found that the majority of Americans believed air pollution has gotten worse in the past 10 years (66%) and would continue to get worse (67%). On a state level, 78% of Californians were reported to believe the state has hardly made progress in solving environmental problems and around 80% of New Yorkers rate air pollution as a “very serious” or “somewhat serious” problem, as do 77% of Texans. Ironically enough, actual ozone level data reflect that the US has made dramatic progress in reducing air pollution over the last few decades, and most American cities now enjoy relatively good air quality. The nation's monitoring stations registered ozone in excess of federal standards more than 12 times per year in the early 1980s, while the end of 2002 saw a fall to 4 exceedances per year. The considerable improvement in air quality should be perceptible, yet there remains a large disconnect between perception and reality. It is likely the presence of confirmation bias prompts individuals to cling to historically negative viewpoints of American air quality and overlook objective improvements, impeding their ability to perceive environmental impacts generated by measures put in place to mitigate air pollution.

An observational study by King (2014) probes into cases where confirmation bias contributes to false perceptions of air pollution due to influence from preconceived notions tied to socioeconomic levels.⁴ The study is set in Chicago, where—in contrast to the

national trend—non-Hispanic Blacks are slightly less likely than other groups to live near toxic-releasing facilities and Hispanics are more likely to be exposed—consistent with a national tendency for higher environmental risk near concentrations of immigrants. Key results are as follows: residents who rated their air quality as “Poor” lived in neighbourhoods where, on average, 21.5% of families were in poverty, conversely, those who rated their air as “Excellent” lived in area with only 10.9% poverty; “Poor” raters also lived in areas that were 37.9% Black and 33.6% Hispanic on average, whereas “Excellent” raters lived in areas that were only 22.5% Black and 20.9% Hispanic; 46.9% of residents with less than a high school education rated their air poorly, while only 28.4% of those with a college degree did the same. Despite the stark perceptual difference, the general difference in objective pollution levels between people who reported “Excellent” air and “Poor” air is actually small. Thus, social cues (in this case, racial demographics) emerge as the key predictors of air quality, a finding consistent with a neighborhood stigma interpretation.

King suggests a “neighbourhood stigma” theory wherein “once a place acquires a negative reputation, respondents will tend to rely heavily on that evaluation when inferring neighborhood characteristics, particularly in the absence of domain-specific information or experience”. While place reputation appears dependent on racial/ethnic phenotypes in Chicago, other considerations may be more salient in other areas. This relates back to confirmation bias by demonstrating how people are inclined to make judgements about air quality that align with preconceived social beliefs.

Fully grasping the implications and applications of confirmation bias can enable policy makers to not only alleviate its negative effects but harness it to positively influence public behaviour. Here I propose two communication strategies regulators can adopt to leverage on confirmation bias in the battle against air pollution.

The first strategy uses confirmation bias to link environmental messaging to identity and values. To do this, policymakers can frame pollution issues in a manner that aligns with different groups’ values. This plays into their confirmation biases and could potentially make them more receptive to proposed environmental actions. For instance, a family health and economic efficiency-oriented narrative can be channelled towards conservatives while climate justice and intergenerational responsibility can be emphasized to progressives. For a conservative audience, messaging could highlight how reducing air pollution protects children's health and lowers long-term healthcare costs. For a progressive audience, there could be a greater focus on the disproportionate impact poor air quality has on marginalised communities.

The second one involves curating source networks and information channels that hold sway over different groups. This approach leverages confirmation bias by strategically delivering the value-aligned messages from the first strategy through sources and channels the target audience already trusts and seeks out. For a conservative audience, this means partnering with and amplifying messages through faith leaders, military veterans' groups, and centre-right media personalities. For a progressive audience, effective channels include social justice organisations, educational institutions, and progressive media outlets.

VI. Loss Aversion

A coin is tossed and the following question is posed to you: "If it's tails, you lose \$10. How much would you have to gain on winning in order for this gamble to be acceptable to you?" Was your answer around \$20? That was the answer most of Kahneman's students gave him. From this demonstration, we can plainly see that people place twice as much value on losing money as opposed to gaining money—a textbook example of the loss aversion hypothesis. (Kahneman, 2011)

The loss aversion hypothesis was first introduced by Kahneman and Tversky (1979) as an essential feature of prospect theory, which suggests people make decisions based on perceived value relative to a reference point instead of a static utility. The hypothesis is that "the response to losses is more extreme than the response to gain. The displeasure associated with losing a sum of money is generally greater than the pleasure associated with winning the same amount" (Kahneman and Tversky, 1981, Pg 454). A consequence of loss aversion is that it can lead to inconsistencies in risk preferences, i.e. the degree to which an individual prefers a sure amount to a gamble for any arbitrary risk. Essentially, an individual is risk-averse if they prefer the sure amount equal to the expected value of the gamble to the gamble, and risk-seeking if they prefer the latter. (Menezes and Hanson, 1970)

Karle et al. (2013) demonstrated the mechanisms of the heuristic in an experimental setting using choices among different lotteries.³² They gave subjects two series of choices, with 6 choices each, where they had to choose between lotteries and sure payments. For Series A, subjects had to make 6 choices between a lottery with 50% chance of winning 1 Euro and 50% chance of winning nothing, and a sure payment of S. S was either 10, 20, 30, 40, 50, or 60 Eurocents. This choice purely elicits their risk preferences without involving loss aversion as it only comprises gains. A risk-neutral individual would be indifferent between the lottery (which has an expected value of 50 Eurocents) and a sure payment of 50 Eurocents. Subjects who choose the lottery over sure payments of 50 Eurocents and above are risk-seekers, while those who choose sure payments of 50 Eurocents and below over

the lottery are risk-averse. In the experiment, 76.7% of participants pick a sure payment that is below or equal to 50 cents (risk-averse).

In series B, subjects had to make six choices between a lottery that gave a $\frac{1}{3}$ chance of winning €1 and a $\frac{2}{3}$ chance of losing R Eurocents, and a sure payment of zero. R was either 0, 10, 20, 30, 50, 70, or 100 Eurocents. This choice is where loss aversion comes into play. Going by the same logic in Series A, a risk-neutral individual would be indifferent between a lottery of expected value zero (when $R=33$) and the sure payment. An individual that, additionally, displays loss aversion would be willing to accept a payment of zero more often for any given R. In the results of the experiment, 16.8% of participants make a choice that is more risk averse compared to what they did in Series A and therefore display a change in behaviour due to loss aversion.

Pollution links closely to loss aversion in terms of environmental action by individuals or groups. Whether or not to undertake a pro-environment action is a choice individuals have to make. That decision making process likely involves weighing private benefit derived against private cost personally incurred. Previous research done on the topic has shown that when faced with two possible choices, one energy-efficient and the other not, people will only choose the energy-efficient item if the payback time for the energy saved is very short (Kollmuss and Agyeman, 2002). This is consistent with loss averse behaviour as they value the loss of their time more than the environmental benefit. Moreover, since partaking in pro-environmental behaviour often requires a lifestyle change and a loss of the familiarity and safety in one's usual routine or incurs implementation costs, loss-averse individuals are likely to opt for the non-environmental option.

Bartczak et al. (2015) explores loss aversion in the environmental domain via a within-subject field study examining choices over a risky environmental good, wildfires in Poland.⁵ The study adapts Tanaka et al.'s (2010) experimental lottery design where Polish individuals are presented with three series of lottery pairs and asked to choose one lottery for each pair.³⁴ Respondents are asked to make a choice between programs that, instead of financial rewards, presented environmental effects of programs to reduce the risk of forest wildfires. In Treatment 1 and 2, they are told that there are two potential programs to further reduce the risk of forest fires which carry different success rates, and results are described as the annual average reduction of burned forest area in Poland and presented as uncertain. Simply put, the first two treatments involve choosing between only gains. As detailed above in Karle et al.'s study, this is used to determine risk preference independent of loss aversion. In Treatment 3 (the core treatment), respondents are asked to choose between programmes which would end in either success or failure, where more forests would be destroyed than currently. As there is now an element of loss, subjects' loss aversion can be elicited. Results from Treatment 1 and 2 indicated 75% of respondents were

risk-averse. However, in Treatment 3 which introduces losses, the majority ($\approx 85\%$) of respondents exhibited risk-seeking behaviour instead. The paper provides conclusive evidence for loss aversion in the environmental context.

More recently, Knobloch et al.'s (2019) paper displays loss aversion's widespread applicability to the global energy model, using residential heating technology uptake as a case study.³⁵ The sector was chosen as decisions are overwhelmingly made by consumers, but not predominantly influenced by other interfering cultural factors. Additionally, the sector accounts for the largest share of direct greenhouse gas emissions by households. It is an observational study that performs analyses at both a global level and the level of individual countries, simulating future heating technology diffusion across 59 world regions.

The study employs a model that simulates consumer decision-making and weighs benefits (convenience, environmental benefits, etc.) and costs (upfront costs, operating costs, etc.). The model is run twice, with default loss aversion parameters and without. The model is used to compare the efficacy of four decarbonization policy scenarios with and without the effect of consumer loss aversion. These policy scenarios involve different levels of residential carbon taxes and subsidy payments. In this case, a carbon tax acts as the "gain" element as it increases the size of relative gains from switching to renewables by reducing energy costs, while subsidies act as the "loss" element as they reduce the size of relative losses. The aim is to assess how loss aversion influences model projections of technological change and policy effectiveness, especially as current energy models often assume optimal consumer decisions, which might lead to overly optimistic expectations of climate policy efficacy. Results indicate that loss-averse decision-makers are less likely to adopt energy-efficient technologies, as higher upfront costs are often perceived as losses with a stronger impact than future energy savings (evaluated as gains). This asymmetric valuation results in a "conservative shift" where preferences for current technologies are strengthened, reducing uptake of newer technologies.

The model further shows that subsidies were relatively more effective than carbon taxes as they directly address the pain point for loss-averse consumers—high investment costs for renewables. On the other hand, projections show that a carbon tax of 200 €/tCO₂ while considering loss aversion and a carbon tax of 100 €/tCO₂ without considering loss aversion achieve a similar reduction in overall emission levels. Thus, taxes require more drastic implementation, which is likely to be received poorly by the public. In short, the model tells us that we need double the tax to see plausible results comparable to idealistic projections (without loss aversion).

These differences in outcomes suggest that consumer behavior, specifically loss aversion, has a profound quantitative impact on the feasibility and required intensity of climate policies.

Having proven how loss aversion pertains to environmental pollution, policymakers can tap on this heuristic to maximise the effectiveness of policies aimed at reducing pollution. While loss aversion is often discussed in the context of how it can hinder the effectiveness of policies, there are now several policies and interventions that have sought to actively “make use of” this behavioral trait to achieve desired outcomes. Current and previous approaches often involve framing choices to emphasize potential losses, thereby leveraging individuals’ stronger aversion to losing something compared to gaining an equivalent amount.

A notable application is in the design of incentives for healthy behaviours, where individuals are offered “insurance” (a sum of money) but told they will lose it if they do not achieve a certain healthy behaviour (e.g. attending an appointment or making a healthy choice). The prospect of losing this already “allocated” amount can be motivating and has been shown to encourage high rates of health behaviour uptake (Meeker et al., 2015).

This strategy can be adapted for pollution reduction in various ways, such as setting household carbon emission thresholds. A “green bonus” can be granted to households at the start of each month (e.g. credit on their utility bill) but lose part or all of it at the end of the month if they exceed the given emission threshold. This creates a stronger motivation to adopt greener appliances and habits than a standard rebate system.

Alternatively, Commuters could begin each week or month with a “streak bonus” (e.g. 10% discount on fares) that diminishes if they fall short of their weekly quota. In other words, missing a day of commuting via public transport would constitute a penalty (e.g. - 0.5% off the discount).

These examples serve to nudge individuals to more willingly adopt greener lifestyle habits.

V. Discussion

Up till this point, each heuristic has been examined independently. However, in practice they do not operate in isolation. Although not limited to these three biases, we see evidence of a perpetual compounding effect even just within the three heuristics examined in this paper, demonstrating a mutually reinforcing relationship not yet explored in previously cited studies. This section proposes a possible account of such an interaction might unfold.

One way of conceptualising this interaction is as a sequential cycle, in which each heuristic builds on the perceptual conditions created by the preceding one. This is not offered as a definitive model but as a plausible account that is broadly consistent with the evidence reviewed in earlier sections.

The halo effect may be understood as the stimulus that establishes the individual's baseline perception of local air quality. The theoretical basis for placing it first in the sequence is that, by definition, the halo effect is a phenomenon of global evaluation—it operates at the level of overall impression, shaping how an individual feels about a place or environment before any deliberate assessment of its specific attributes has taken place (Nisbett and Wilson, 1977). As demonstrated by Hofflinger et al. (2019) and Boso et al. (2020), (positive) emotional attachment to one's home and neighbourhood generates a global positive evaluation that colours assessment of individual attributes, including air quality, in a more favourable direction than objective data warrants. This initial perceptual anchor is not consciously constructed; it emerges automatically from the same System 1 processing that underlies heuristic reasoning more broadly. The critical point is that once established, this baseline does not simply sit passively in the individual's mind, it actively colours how subsequent information is processed.

This is where the confirmation bias plausibly enters. Having formed an overly optimistic baseline perception, the individual is now predisposed to attend selectively to information that confirms it. On a clear, sunny day, the absence of visible haze may be registered as evidence of clean air; a health advisory may be discounted as overstated; a neighbour's complaint about air quality may be attributed to hypersensitivity. As Schwartz (2004) and King (2014) demonstrate, such selective interpretation is capable of sustaining misperceptions across years and even decades, insulating individuals from the corrective influence of objective environmental measures. In this case, the halo-derived baseline provides the very belief that the confirmation bias then works to protect and entrench.

The third stage introduces loss aversion as a barrier to behavioural correction. Even in cases where the first two stages are partially overcome—where an individual does come to accept that local air quality is worse than they had assumed—the translation of updated perception into action faces a further obstacle. As Knobloch et al. (2019) and Bartczak et al., (2015) elucidate, the upfront costs, lifestyle disruptions, and sacrifices of familiarity associated with pro-environmental behaviour are experienced as losses, weighted disproportionately against the future gains of cleaner air. Loss aversion, understood as such, functions as the final barrier in the cycle. It is the mechanism by which even a corrected perception fails to produce a corrected behaviour. The individual may acknowledge the problem and yet remain inert.

Taken together, this proposed sequence suggests a possible dynamic in which the halo effect generates the misconception, confirmation bias reinforces it, and loss aversion prevents actionable progress even when negative perceptions are dislodged. While the precise causal ordering proposed above remains an open empirical question, the critical takeaway is that when multiple heuristics operate simultaneously in the same cognitive environment, the result may be qualitatively different from, and more intractable than the basic sum of their individual effects.

Two of the cases discussed in earlier sections are worth revisiting through this integrative lens, as they are suggestive—if not conclusive—of the kind of compounding dynamic described above.

The first is the wood-burning stove context in Southern Chile examined by Boso et al. (2020). Residents in Temuco and Padre Las Casa systemically overestimated indoor air quality relative to both sensor measurements and inspector evaluations—a pattern consistent with the home halo effect. Yet this misperception persisted despite aggressive government intervention, including refurbishment subsidies and prohibitions on firewood use, suggesting that objective information and material incentives that should rationally tip the scale in favour of actionable change were insufficient to shift behaviour. This resistance is arguably consistent with confirmation bias entrenching the favourable baseline perception against contradictory policy messaging, while loss aversion simultaneously raised the psychological cost of abandoning familiar heating practices.

The second case is King's (2014) study of air quality perception in Chicago. This illustrates that the compounding dynamic need not operate exclusively in the direction of over-optimism. Here, neighbourhood stigma (negative halo effect), generated a pessimistic baseline perception of air quality in predominantly Black and Hispanic neighbourhoods, despite the objective difference in pollution levels between high- and low-rated areas being small. Confirmation bias then plausibly entrenched this stigma, causing residents to interpret ambiguous environmental cues through a lens of pre-existing negative expectations. Loss aversion may further have discouraged investment in local environmental improvement, with residents perceiving the costs of action as disproportionate given the low perceived likelihood of meaningful change. The cycle, in this case, runs in the opposite direction—producing systematic underestimation of air quality rather than overestimation—but the underlying compounding logic is arguably the same.

These two cases do not constitute proof of the proposed interaction. They are better understood as illustrations of contexts in which such an interaction is a plausible

contributing explanation for outcomes that individual heuristics alone do not fully account for.

If the compounding dynamic proposed here has merit, it carries meaningful implications for the design of pollution-related policy. Interventions targeting only one heuristic at a time may find their effectiveness partially undermined by the others operating in the background. For example, a communication campaign that successfully challenges a community's confirmation bias, for instance, may still fail to produce behavioural change if loss aversion remains unaddressed—the individual may update their beliefs without updating their behaviour.

This suggests that layered policy approaches—ones that address the emotional baseline, the information environment, and the incentives structure in coordinated fashion—may be more effective than single-lever interventions, even if they are considerably more complex to design and evaluate. The individual policy proposals outlined in Sections II, III, and IV might therefore be understood not merely as standalone recommendations but as potentially complementary components of a broader behavioural strategy but as a potentially complementary component of a broader behavioural strategy: reshaping the halo first to establish a more accurate perceptual baseline, then delivering value-aligned messaging to work with rather than against confirmation bias.

VI. Conclusion

This review has provided compelling evidence that heuristics—namely the halo effect, confirmation bias, and loss aversion—profoundly shape individual and collective perceptions of air quality, often creating a significant gap between objective reality and subjective interpretation. Findings explored in this paper are not merely curiosities of psychology but demonstrations of powerful, systematic forces that undermine the effectiveness of conventional pollution mitigation policies which assume rational, information-driven public behaviour.

The halo effect explains how positive emotional attachments to one's living environment as a whole colours objective assessment of individual traits like air quality, leading to consistent overestimations of air quality that could create a false sense of security and insulate people from the urgency of environmental risks. Confirmation bias leads communities to interpret ambiguous environmental cues through the lens of pre-existing beliefs and social stigmas, causing real improvements in air quality to have less than proportionate social impacts, and may even lead to ill-targeted policy actions. Loss

aversion reveals a fundamental asymmetry in decision-making: the perceived pain of sacrificing convenience, money, or routine to adopt green behaviours far outweighs the perceived gain of future environmental benefits, creating a powerful inertia against pro-environmental action.

Section V tentatively proposes that these heuristics may do more than operate in parallel. If the compounding dynamic there has merit, the picture that emerges is one in which the halo effect establishes a distorted perceptual baseline, confirmation bias entrenches it and protects the baseline from corrective information, and loss aversion prevents updates perception from translating into behavioural change. The result may be a cycle of misperception that is qualitatively more resistant to intervention than any individual bias would suggest. This interaction remains empirically underexplored, and its precise mechanisms and casual orderings are open questions that future research would do well to address, ideally through experimental and field-based studies designed specifically to examine how multiple heuristics operate simultaneously within environmental decision-making contexts.

Crucially, these heuristics not only present obstacles but also offer novel pathways for policy innovation. Effective countervailing policies work with, not against these ingrained psychological patterns. As proposed, this can involve the following. We can leverage the halo effect by associating clean air technologies with socially desirable status and community identity. We can harness confirmation bias by framing environmental messages to align with deeply held values of different demographic groups, making sustainable choices feel like affirmations of identity instead of contradicting beliefs. Lastly, we can draw on people's loss aversion by restructuring incentives around the avoidance of loss—such as bonds forfeited for non-compliance with green policies. The interaction framework further suggests that these interventions may be most effective not as standalone measures but as coordinated components of a layered behavioural strategy.

Bridging the gap between air quality data and public perception to beget real, effective action requires a behavioural turn in environmental policy. We must go beyond mere information dissemination to strategically reshaping the cognitive and emotional landscape in which decisions are made. By integrating insights from behavioural science into policy design, regulators can craft more resonant interventions that align public perception with planetary health, ultimately transforming heuristics from barriers into catalysts for a cleaner future.

Bibliography

- (1) World Health Organization: WHO. (2019, July 30). Air pollution. https://www.who.int/health-topics/air-pollution#tab=tab_2
- (2) Kahneman, D., Smith, V. (2002) Advanced Information on the Prize in Economic Sciences. <https://www.nobelprize.org/uploads/2018/06/advanced-economicsciences2002-1.pdf>.
- (3) Boso, À., Álvarez, B., Oltra, C., Garrido, J., Muñoz, C., & Galvez-García, G. (2020). The grass is always greener on my side: A field experiment examining the home halo effect. *Sustainability*, 12(16), 6335. <https://doi.org/10.3390/su12166335>.
- (4) King, K. E. (2014). Chicago Residents' Perceptions of Air Quality: Objective Pollution, the Built Environment, and Neighborhood Stigma Theory. *Population and Environment*, 37(1), 1–21. <https://doi.org/10.1007/s11111-014-0228-x>
- (5) Bartczak, A., Chilton, S., & Meyerhoff, J. (2015). Wildfires in Poland: The impact of risk preferences and loss aversion on environmental choices. *Ecological Economics*, 116, 300–309. <https://doi.org/10.1016/j.ecolecon.2015.05.006>
- (6) Henning, R. J. (2024). Particulate Matter Air Pollution is a Significant Risk Factor for Cardiovascular Disease. *Current Problems in Cardiology*, 49(1, Part B), 102094. <https://doi.org/10.1016/j.cpcardiol.2023.102094>
- (7) UNEP. (2024). How Asian countries could save lives, boost growth by tackling air pollution. UNEP. <https://www.unep.org/news-and-stories/story/how-asian-countries-could-save-lives-boost-growth-tackling-air-pollution>
- (8) Institute for Health Metrics and Evaluation: IHME. (2023). Global Burden of Disease 2023: Findings from the GBD 2023 study. Institute for Health Metrics and Evaluation. <https://www.healthdata.org/research-analysis/library/global-burden-disease-2023-findings-gbd-2023-study>
- (9) Thanakon Sukuman, Ueda, K., Sarunya Sujaritpong, Hirunwut Praekunatham, Kornwipa Punnasiri, Tuangsit Wimuktayon, & Trakarn Prapaspongsa. (2023). Health

Impacts from PM2.5 Exposure Using Environmental Epidemiology and Health Risk Assessment: A Review. *Applied Environmental Research*, 1–14.
<https://doi.org/10.35762/aer.2023010>

- (10) Myllyvirta, L. (2020). Quantifying the Economic Costs of Air Pollution from Fossil Fuels Key messages. Retrieved February 15, 2026, from
<https://energyandcleanair.org/wp/wp-content/uploads/2020/02/Cost-of-fossil-fuels-briefing.pdf>.
- (11) Royal College of Physicians: RCP. (2025). A breath of fresh air: responding to the health challenges of modern air pollution. Rcp.ac.uk.
<https://www.rcp.ac.uk/policy-and-campaigns/policy-documents/a-breath-of-fresh-air-responding-to-the-health-challenges-of-modern-air-pollution/>
- (12) CBI Economics. (2020). Breathing life into the UK economy: Quantifying the economic benefits of cleaner air. (2020).
https://s40026.pcdn.co/wp-content/uploads/Breathing_life_into_the_UK_economy-Report-Final.pdf.
- (13) Berg, L. (2021, September). More than a Public Health Crisis: Air Pollution's Impact on the Indian Economy. Dalberg.
<https://dalberg.com/our-ideas/more-than-a-public-health-crisis-air-pollutions-impact-on-the-indian-economy/>
- (14) Fang, T., Gu, Y., Li, J., Sung, J. J. Y., Yim, S. H. L. (2025). Effects of Climate Change on PM2.5 Formation Chemistry, and the Associated Human Health Impacts and Economic Costs in Southeast Asia. *Environment International*, 205, 109835.
<https://doi.org/10.1016/j.envint.2025.109835>
- (15) Kahneman, D. (2011). *Thinking, Fast and Slow*. Penguin.
- (16) Nisbett, R. E., & Wilson, T. D. (1977). The Halo Effect: Evidence for Unconscious Alteration of Judgments. *Journal of Personality and Social Psychology*, 35(4), 250–256.
<https://doi.org/10.1037/0022-3514.35.4.250>
- (17) Hofflinger, Á., Boso, À., & Oltra, C. (2019). The Home Halo Effect: how Air Quality Perception is Influenced by Place Attachment. *Human Ecology*, 47(4), 589–600.
<https://doi.org/10.1007/s10745-019-00100-z>

- (18) D'Antoni, D., Smith, L., Auyeung, V., & Weinman, J. (2017). Psychosocial and Demographic Predictors of Adherence and Non-adherence to Health Advice Accompanying Air Quality Warning Systems: A Systematic Review. *Environmental Health*, 16(1). <https://doi.org/10.1186/s12940-017-0307-4>
- (19) Bickerstaff, K., & Walker, G. (2001). Public understandings of air pollution: the “localisation” of environmental risk. *Global Environmental Change*, 11(2), 133–145. [https://doi.org/10.1016/S0959-3780\(00\)00063-7](https://doi.org/10.1016/S0959-3780(00)00063-7)
- (20) Rankin, R. E. (1969). Air Pollution Control and Public Apathy. *Journal of the Air Pollution Control Association*, 19(8), 565–569. <https://doi.org/10.1080/00022470.1969.10466523>
- (21) Nickerson, R. S. (1998). Confirmation bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- (22) Popper, K. R. (1977). The Logic of Scientific Discovery. *Systematic Zoology*, 26(3), 361. <https://doi.org/10.2307/2412687>
- (23) Wason, P. (1960). On the Failure to Eliminate Hypotheses in a Conceptual Task. *Quarterly Journal of Experimental Psychology*, 12(3), 129–140. <https://doi.org/10.1080/17470216008416717>
- (24) Oswald, M. E., Grosjean, S. (2012). Confirmation bias. In R. F. Pohl (Eds.), *Cognitive Illusions. A Handbook on Fallacies and Biases in Thinking, Judgement and Memory* (pp 91–108). Psychology Press.
- (25) Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098–2109.
- (26) Fowler, D. (2020). A Chronology of Global Air Quality. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 378(2183). <https://doi.org/10.1098/rsta.2019.0314>

- (27) Schwartz, J. (2024). *Clearing the Air*. Ssrn.com.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=505604
- (28) Ramírez, A. S., Ramondt, S., Van Bogart, K., & Perez-Zuniga, R. (2019). Public Awareness of Air Pollution and Health Threats: Challenges and Opportunities for Communication Strategies To Improve Environmental Health Literacy. *Journal of Health Communication*, 24(1), 75–83. <https://doi.org/10.1080/10810730.2019.1574320>
- (29) Kahneman, D., & Tversky, A. (1979). Prospect theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2), 263–292. <https://doi.org/10.2307/1914185>
- (30) Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- (31) Menezes, C. F., & Hanson, D. L. (1970). On the Theory of Risk Aversion. *International Economic Review*, 11(3), 481. <https://doi.org/10.2307/2525326>
- (32) Karle, H., Kirchsteiger, G., & Peitz, M. (2015). Loss Aversion and Consumption Choice: Theory and Experimental Evidence. *American Economic Journal: Microeconomics*, 7(2), 101–120. <https://doi.org/10.1257/mic.20130104>
- (33) Kollmuss, A., & Agyeman, J. (2002). Mind the Gap: Why Do People Act Environmentally and What Are the Barriers to pro-environmental behavior? *Environmental Education Research*, 8(3), 239–260.
<https://doi.org/10.1080/13504620220145401>
- (34) Tanaka, T., Camerer, C. F., & Nguyen, Q. (2006). Risk and Time Preferences: Experimental and Household Survey Data from Vietnam. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.877229>

- (35) Knobloch, F., Huijbregts, M. A. J., & Mercure, J.-F. (2019). Modelling the effectiveness of climate policies: How important is loss aversion by consumers? *Renewable and Sustainable Energy Reviews*, 116, 109419. <https://doi.org/10.1016/j.rser.2019.109419>
- (36) Meeker, D., Thompson, C., Strylewicz, G., Knight, T. K., & Doctor, J. N. (2015). Use of Insurance Against a Small Loss as an Incentive Strategy. *Decision Analysis*, 12(3), 122–129. <https://doi.org/10.1287/deca.2015.0314>

Thank you for your thoughtful comments.

I have carefully considered each comment and made substantive revisions in response. Below, I address each concern in turn. Your given comments are in inverted commas, followed by my responses. Parts in *italics* are quotations from the revised version of the manuscript that respond to the various comments.

1. Methodology and Study Selection

- a. Comment: "The most significant limitation of the paper is the absence of a clearly defined and transparent methodology... The author(s) should clarify whether the approach follows a systematic, scoping, or narrative review framework and justify the chosen method."
- b. Response: Section II (pages 4-5) in the revised manuscript is a completely new methodology section. The paper adopts a narrative review framework, selected on the grounds that its primary objective is thematic synthesis and theoretical integration across behavioural economics, environmental science, and public policy — rather than an exhaustive enumeration of available studies. This choice is now explicitly justified in the manuscript:

II. Methodology

This paper adopts a narrative review framework to synthesise existing empirical and theoretical literature on the role of cognitive heuristics in shaping public perception of air quality. A narrative approach was selected on the grounds that the primary objective of this review is thematic synthesis and theoretical integration across disciplines, namely behavioural economics, environmental science, and public policy, rather than an exhaustive enumeration of all available studies.

Literature was identified through searches of Google Scholar, PubMed, Web of Science, and JSTOR, supplemented by reports from international organisations including the World Organisation (WHO), the United Nations Environment Programme (UNEP), and the Institute for Health Metrics and Evaluation (IHME). Search terms included combinations of the following: "air pollution", "halo effect", "confirmation bias", "loss aversion", "air quality perception", and "cognitive bias". No strict date restriction was imposed, as the review deliberately includes foundational experimental studies and theoretical papers, such as Kahneman and Tversky's on prospect theory (1979) and Wason's on the rule discovery task

(1960), providing a core, well-tested psychological base to build upon. More recent empirical studies, prioritised from 2000 onwards, were selected to ground these psychological frameworks in contemporary environmental contexts.

Studies were included based on relevance to at least one of the three heuristics under examination, methodological transparency and publication in peer-reviewed journals or by credible international institutions; purely anecdotal or non-empirical commentary was excluded. Findings were then organised thematically; for each heuristic, foundational experimental evidence establishing the psychological mechanism was identified first, followed by domain-specific studies applying that mechanism to air pollution perception or responsive behaviour. Policy implications were derived inductively from the patterns observed across studies within each thematic cluster, with the aim of translating behavioural insights into actionable regulatory recommendations.

2. Introduction: research gap and novelty

- a. Comment: "The introduction should more explicitly highlight what is missing in current research, how this paper responds to that gap, and why this contribution is novel... the value added by this review is not sufficiently clear."
- b. Response: The introduction has been revised to foreground the paper's contribution more directly. In particular, the following paragraph on page 3 explicitly identifies the gap and shows why this paper is novel:

There exists a growing body of literature on both air pollution and behavioural economics individually, but a critical gap persists at their intersection. Where behavioural approaches have been recently applied to environmental risk perception, they tend to examine cognitive biases within a narrow empirical scope, rather than developing an integrated account from various contexts of how the general bias impacts air quality perception as a whole. The systemic application of heuristics—namely the halo effect confirmation bias, and loss aversion—to air quality perception and pollution-related policy design therefore remains underdeveloped in existing literature. This paper responds to that gap by examining each of these three heuristics in turn, synthesising psychological and environmental evidence to demonstrate how each independently distorts public understanding of air quality and undermines conventional policy responses. The

paper's key contribution lies in moving beyond the diagnosis of cognitive bias to the derivation of targeted, actionable policy recommendations that work with these psychological tendencies.

3. Results and Discussion: integration and analytical presentation

- a. Comment: "The results and discussion sections require stronger integration and a more analytical presentation... findings should be systematically organised, potentially through thematic synthesis or analytical categorisation."
- b. Response: A new Discussion section, Section V, has been added that integrates findings across all three heuristics, proposing a sequential compounding model in which the halo effect establishes a misperceived baseline, confirmation bias entrenches it, and loss aversion prevents behavioural correction even when the misperception is partially overcome. This section explicitly connects the paper's analytical findings to existing behavioural theory and policy contexts. Language and structure have also been reviewed and tightened across the manuscript. The new section is as follows:

V. *Discussion*

Up till this point, each heuristic has been examined independently. However, in practice they do not operate in isolation. Although not limited to these three biases, we see evidence of a perpetual compounding effect even just within the three heuristics examined in this paper, demonstrating a mutually reinforcing relationship not yet explored in previously cited studies. This section proposes a possible account of such an interaction might unfold.

One way of conceptualising this interaction is as a sequential cycle, in which each heuristic builds on the perceptual conditions created by the preceding one. This is not offered as a definitive model but as a plausible account that is broadly consistent with the evidence reviewed in earlier sections.

The halo effect may be understood as the stimulus that establishes the individual's baseline perception of local air quality. The theoretical basis for placing it first in the sequence is that, by definition, the halo effect is a phenomenon of global evaluation—it operates at the level of overall impression, shaping how an individual feels about a place or environment before any deliberate assessment of its specific attributes has taken place (Nisbett and Wilson, 1977). As demonstrated by Hofflinger et al. (2019) and Boso et al.

(2020), (positive) emotional attachment to one's home and neighbourhood generates a global positive evaluation that colours assessment of individual attributes, including air quality, in a more favourable direction than objective data warrants. This initial perceptual anchor is not consciously constructed; it emerges automatically from the same System 1 processing that underlies heuristic reasoning more broadly. The critical point is that once established, this baseline does not simply sit passively in the individual's mind, it actively colours how subsequent information is processed.

This is where the confirmation bias plausibly enters. Having formed an overly optimistic baseline perception, the individual is now predisposed to attend selectively to information that confirms it. On a clear, sunny day, the absence of visible haze may be registered as evidence of clean air; a health advisory may be discounted as overstated; a neighbour's complaint about air quality may be attributed to hypersensitivity. As Schwartz (2004) and King (2014) demonstrate, such selective interpretation is capable of sustaining misperceptions across years and even decades, insulating individuals from the corrective influence of objective environmental measures. In this case, the halo-derived baseline provides the very belief that the confirmation bias then works to protect and entrench.

The third stage introduces loss aversion as a barrier to behavioural correction. Even in cases where the first two stages are partially overcome—where an individual does come to accept that local air quality is worse than they had assumed—the translation of updated perception into action faces a further obstacle. As Knobloch et al. (2019) and Bartczak et al., (2015) elucidate, the upfront costs, lifestyle disruptions, and sacrifices of familiarity associated with pro-environmental behaviour are experienced as losses, weighted disproportionately against the future gains of cleaner air. Loss aversion, understood as such, functions as the final barrier in the cycle. It is the mechanism by which even a corrected perception fails to produce a corrected behaviour. The individual may acknowledge the problem and yet remain inert.

Taken together, this proposed sequence suggests a possible dynamic in which the halo effect generates the misconception, confirmation bias reinforces it, and loss aversion prevents actionable progress even when negative perceptions are dislodged. While the precise causal ordering proposed above remains an open

empirical question, the critical takeaway is that when multiple heuristics operate simultaneously in the same cognitive environment, the result may be qualitatively different from, and more intractable than the basic sum of their individual effects.

Two of the cases discussed in earlier sections are worth revisiting through this integrative lens, as they are suggestive—if not conclusive—of the kind of compounding dynamic described above.

The first is the wood-burning stove context in Southern Chile examined by Boso et al. (2020). Residents in Temuco and Padre Las Casa systemically overestimated indoor air quality relative to both sensor measurements and inspector evaluations—a pattern consistent with the home halo effect. Yet this misperception persisted despite aggressive government intervention, including refurbishment subsidies and prohibitions on firewood use, suggesting that objective information and material incentives that should rationally tip the scale in favour of actionable change were insufficient to shift behaviour. This resistance is arguably consistent with confirmation bias entrenching the favourable baseline perception against contradictory policy messaging, while loss aversion simultaneously raised the psychological cost of abandoning familiar heating practices.

The second case is King's (2014) study of air quality perception in Chicago. This illustrates that the compounding dynamic need not operate exclusively in the direction of over-optimism. Here, neighbourhood stigma (negative halo effect), generated a pessimistic baseline perception of air quality in predominantly Black and Hispanic neighbourhoods, despite the objective difference in pollution levels between high- and low-rated areas being small. Confirmation bias then plausibly entrenched this stigma, causing residents to interpret ambiguous environmental cues through a lens of pre-existing negative expectations. Loss aversion may further have discouraged investment in local environmental improvement, with residents perceiving the costs of action as disproportionate given the low perceived likelihood of meaningful change. The cycle, in this case, runs in the opposite direction—producing systematic underestimation of air quality rather than overestimation—but the underlying compounding logic is arguably the same.

These two cases do not constitute proof of the proposed interaction. They are better understood as illustrations of contexts in which such an interaction is a plausible contributing explanation for outcomes that individual heuristics alone do not fully account for.

If the compounding dynamic proposed here has merit, it carries meaningful implications for the design of pollution-related policy. Interventions targeting only one heuristic at a time may find their effectiveness partially undermined by the others operating in the background. For example, a communication campaign that successfully challenges a community's confirmation bias, for instance, may still fail to produce behavioural change if loss aversion remains unaddressed—the individual may update their beliefs without updating their behaviour.

This suggests that layered policy approaches—ones that address the emotional baseline, the information environment, and the incentives structure in coordinated fashion—may be more effective than single-lever interventions, even if they are considerably more complex to design and evaluate. The individual policy proposals outlined in Sections II, III, and IV might therefore be understood not merely as standalone recommendations but as potentially complementary components of a broader behavioural strategy but as a potentially complementary component of a broader behavioural strategy: reshaping the halo first to establish a more accurate perceptual baseline, then delivering value-aligned messaging to work with rather than against confirmation bias.

I hope the revised manuscript attached addresses the reviewers' concerns satisfactorily.

The Psychology of Pollution: Analyzing How Heuristics Affect Perception of Air Quality and Policy Design

The manuscript presents an ambitious and relevant interdisciplinary review that brings together behavioral economics, environmental risk perception, and policy design. The comments below are organized by section and, where possible, anchored to section names and opening phrases to help the author locate each point as easily as possible. The review focuses on argument development, use of evidence, scientific rigor, organization, grammar, and academic writing style. The goal is to provide constructive feedback that strengthens the manuscript while preserving the author's ideas and voice..

Title, Keywords, and Abstract

- Title: The title is engaging and accurately signals the interdisciplinary focus. Consider whether “Analysing” and “Behavioral” should follow one spelling convention throughout the manuscript, since British and American English are currently mixed.
- Keywords: The keywords are relevant, but “Policy Design” should follow the same capitalization style as the other terms. You may also want to consider whether “risk perception” is central enough to include.
- Abstract, opening sentence: The opening establishes significance quickly, but “severe, global issue” is slightly general. Consider whether the sentence could identify air pollution as an environmental health risk rather than relying on broad emphasis.
- Abstract, sentence beginning “Heuristics”: The phrase “due its subjectivity and complexity” needs a grammatical correction and a clearer explanation of what is subjective. Air pollution itself is measurable, while public perception of it is subjective.
- Abstract, sentence beginning “This paper conducts an analysis”: “Conducts an analysis into” is awkward. The sentence would read more clearly if it stated that the paper examines the role of the three heuristics.
- Abstract, halo effect sentence: The claim that the halo effect “causes” residents to overestimate air cleanliness is stronger than the reviewed evidence appears to support. Consider using language that reflects association or contribution unless the cited studies establish causation.
- Abstract, confirmation bias sentence: This is clear, but it presents one direction of the bias only. The paper later discusses both optimistic and pessimistic distortions. Consider whether the abstract should reflect that broader application.
- Abstract, loss aversion sentence: The connection to carbon taxes is interesting, but “policies that trigger individuals’ loss aversion” is broad. It may help to specify that the issue concerns how costs and incentives are framed and perceived.
- Abstract, sentence beginning “Through a review”: The abstract would be stronger if it identified the paper as a narrative review, since the methodology section makes that distinction important.
- Abstract, final sentence: “Mitigate and/or strategically harness” is cumbersome. Consider choosing one clear construction and stating that the policies either reduce the effects of these biases or use them to improve policy response.
- Abstract, overall: The abstract captures the main argument well, but it would benefit from a brief indication that the proposed interaction among the three heuristics is conceptual and remains to be tested empirically.

I. Introduction

- Introduction, opening paragraph, sentence beginning “Air pollution stands”: “Environmental causes” does not seem to be the intended term. Consider whether “environmental threats” or “environmental health risks” better expresses the point.
- Introduction, opening paragraph, economic-cost sentence: The sentence is a fragment. It also uses highly emphatic wording. Consider integrating the economic cost into the preceding or following sentence and maintaining a more measured academic tone.

- Introduction, opening paragraph, sentence beginning “Yet, societal and individual responses”: This is a strong transition from the physical problem to the psychological argument. The phrase “singular threat” may be unnecessary because the paper later acknowledges that pollution risk varies across settings.
- Introduction, behavioral economics paragraph: The contrast between neoclassical and behavioral economics is useful, but it is presented as more absolute than most economic theory would support. Consider acknowledging that neoclassical models often simplify behavior rather than literally assuming that all actors possess perfect information.
- Introduction, behavioral economics paragraph, final citation: The citation is placed as a separate sentence. It should be attached to the claim it supports.
- Introduction, sentence beginning “This paper argues”: The thesis is clear, but “we must” gives it a persuasive tone. A more neutral formulation may be more appropriate for a review article.
- Introduction, paragraph beginning “The imperative for this behavioral approach”: The sentence stating that the most harmful pollutants are imperceptible is too broad. Some hazardous pollutants can produce visible or sensory cues, while PM2.5 itself is generally not directly visible. Consider narrowing the claim.
- Introduction, same paragraph, phrase “human sense”: This should be plural or otherwise clarified because the paragraph later refers to sight, smell, and physical discomfort.
- Introduction, same paragraph: The examples of visual haze, smell, media reports, and discomfort are useful. Consider distinguishing direct sensory cues from socially supplied information, since they may activate different judgment processes.
- Introduction, paragraph beginning “Understanding these cognitive processes”: This effectively explains the policy relevance. “Instrumental data” is unclear in this context. “Objective measurements” or a similarly specific term may be easier to understand.
- Introduction, literature-gap paragraph: The paper identifies a meaningful gap, but the claim that the integrated application “remains underdeveloped” needs stronger support. Consider citing one or two prior reviews or explicitly noting the scope of the search that led to this conclusion.
- Introduction, literature-gap paragraph, list of heuristics: A comma is missing between “halo effect” and “confirmation bias.” There is also a stray “z” at the end of the paragraph.
- Introduction, roadmap (spanning sections) paragraphs: The manuscript explains the contents several times. These paragraphs could be reduced substantially. One concise roadmap would be enough and would create more space for argument development.
- Methodology, roadmap paragraph beginning “The paper proceeds as follows”: The section numbering does not match the manuscript. Methodology is Section II, Background is Section III, and the heuristic sections begin with Section IV. Please correct all section references after the final structure is set.
- Methodology, sentence beginning “The discussion of loss aversion”: “Perceptive losses” is not the right term. Consider wording that refers to perceived losses, upfront costs, or lifestyle sacrifices.
- Introduction, overall: The introduction makes a strong case for the topic, but it would benefit from a clearer separation among three elements: the public health problem, the behavioral theory, and the specific contribution of this review. At present, these elements overlap and occasionally repeat one another.

II. Methodology

- Methodology, opening paragraph: The narrative review design is appropriate for the interdisciplinary aim, and the rationale is explained well. Consider adding a brief statement that the review is interpretive rather than systematic so readers understand the limits of reproducibility.

- Methodology, search strategy: The databases and organizations are listed clearly. The methodology would be stronger if the author reported when the searches were conducted and whether the same search combinations were used across databases.
- Methodology, organization name: “World Organisation” appears to be incomplete. It should identify the World Health Organization accurately. Please be sure to not alternate between US and British Spellings. Keep with what the manuscript allows.
- Methodology, search terms: The listed terms cover the core concepts, but policy-related terms are absent even though policy design is central to the title and conclusion. Consider whether the search strategy included terms such as environmental policy, risk communication, incentives, or behavioral intervention.
- Background, sentence beginning “No strict date restriction”: The rationale for including foundational work is sound. The wording around “Kahneman and Tversky’s on prospect theory” and “Wason’s on the rule discovery task” is incomplete and needs grammatical correction.
- Background, inclusion criteria: The criteria are reasonable but remain broad. Consider clarifying how credibility was judged for institutional reports and how methodological transparency was assessed.
- Background, exclusion criteria: “Purely anecdotal or non-empirical commentary” is clear, but the manuscript later makes policy proposals that are not directly tested. It would help to distinguish literature-selection criteria from the author’s own conceptual recommendations.
- Background, thematic organization: This is one of the strongest parts of the methodology because it explains how foundational studies and environmental applications were linked. Consider making the same sequence visible in each substantive section through consistent subheadings.
- Background, policy implications sentence: “Derived inductively” is a meaningful methodological claim. Consider explaining how the author moved from the evidence to the proposed recommendations, especially where a policy idea is original rather than directly supported by a reviewed intervention study.
- Methodology, overall: The section is a good start, but the review would be easier to evaluate if it included the approximate number of sources screened and retained, even if a formal systematic-review flow diagram is not required.

III. Background

- Background, opening paragraph: This paragraph clearly establishes the purpose of the section. “Personalised human interpretation of objective reality” is somewhat abstract. Consider using more direct wording about how measured pollution levels are interpreted by individuals.
- Background, definition of air pollution: The definition is appropriate. Consider whether quotation marks and citation style are consistent with the rest of the manuscript.
- Background/Methodology, pollutant overview: This is useful background, but the paper later focuses mainly on PM2.5. Consider explaining earlier why PM2.5 is the primary example while still acknowledging that different pollutants are perceived differently.
- Background, global burden paragraph: The health statistics establish significance, but several figures are presented in rapid succession. Consider grouping them by global burden, regional burden, and disease outcomes so the reader can follow the evidence more easily.
- Background, sentence beginning “As a whole, air pollution accounts”: “The culprit of” is informal and grammatically awkward. The relationship between the 1-in-8 estimate and the 7.9 million figure should also be checked for consistency and source year.
- Background, sentence beginning “In particular, studies have shown”: The list of diseases is broad. Consider distinguishing conditions with well-established causal evidence from those for which the association is still emerging.

- Background, economic-cost paragraph: The range of international examples adds breadth, but the paragraph is overloaded. Consider separating global estimates from country examples and limiting the number of statistics to those most relevant to the argument.
- Background, India sentence beginning “Had the country managed”: The grammar is incorrect, and the counterfactual claim needs especially careful wording. Check the original source to make sure the time period and estimated GDP effect are represented accurately.
- Background, sentence beginning “In Southeast Asia”: “Snowballing to a hefty” is journalistic. A more neutral expression would fit the rest of the review.
- Background, closing sentence “Undoubtedly”: Avoid absolute or rhetorical language. The evidence supports a strong economic connection without needing this qualifier.
- Background, paragraph beginning “While the physical and sociological facts”: “Sociological facts” is not clearly defined. Consider whether “health and economic evidence” better reflects the preceding discussion.
- Background, PM2.5 description: The sentence is grammatically difficult and could mislead readers into thinking PM2.5 is always invisible in every context. Consider clarifying that individual particles are not visible and that ambient haze is an imperfect proxy.
- Background, sentence ending with the Fang citation: The citation is separated from the sentence by a period. Please standardize citation placement throughout.
- Background/Halo Effect, System 1 and System 2 paragraph: The framework is relevant, but the explanation is repetitive and includes a stray superscript-style “15.” The sentence beginning “This operation” does not have a clear referent.
- Halo Effect, sentence beginning “Heuristics are the rules of thumb”: This is clear, but “generally” is repeated. Consider revising for smoother flow.
- Halo Effect, cue-substitution paragraph: The examples are useful and directly support the paper’s thesis. The question mark in “Does it smell bad?” should not be followed by a period.
- Halo Effect, same paragraph: The claim that physical cues create a shift from “data-rich” to “cue-driven” decision making is plausible but would benefit from a direct citation.
- Halo Effect, final sentence: “I further probe into” is informal and shifts into first person. Consider a more neutral transition.
- Background, overall: The section successfully establishes why air pollution is difficult to perceive. It would be stronger if the health and economic statistics were condensed and more space were given to the specific psychological problem of judging invisible or delayed risks.

IV. Halo Effect

- Halo Effect, opening example: The everyday example is accessible, but the second-person framing is conversational. Consider introducing the concept in a more academic voice while keeping the example brief.
- Halo Effect, sentence beginning “This is due to the human tendency”: “Base our general judgement off” is informal. The statement also implies that one impression always determines all later judgments. Consider more cautious wording.
- Halo Effect, sentence beginning “The halo effect was first coined”: The effect was named or described by Thorndike, not “first coined by” the person. Check the historical phrasing.
- Halo Effect, quoted definition: Confirm the page number and citation style. The manuscript alternates among “Pg,” parenthetical author-date citations, and numbered references.
- Halo Effect, Nisbett and Wilson experiment: The study is explained clearly, but the amount of procedural detail may be more than the review needs. Consider reducing the description and using the

saved space to explain the limitations of applying a person-perception experiment to place-based air-quality judgments.

- Halo Effect, figure caption: Standardize figure-caption formatting and verify permission or fair-use requirements for reproducing the figure. The image is low resolution and may not reproduce well in publication.
- Halo Effect, interpretation paragraph: Verb tense shifts from “rated” to “rate.” Please keep the description in past tense.
- Halo Effect, sentence beginning “The results quantitatively prove”: “Prove” is too strong for a single experiment. Consider stating that the results support or demonstrate evidence of the halo effect under the study conditions.
- Halo Effect, paragraph beginning “Fundamentally”: This makes the transition to air quality effectively. The phrase “global evaluations are powerful enough to trump other objective data” is stronger than the cited evidence supports and should be qualified.
- Halo Effect, sentence beginning “As a result, vigorous public awareness campaigns”: The sentence makes a large causal leap from the halo effect to persistent mortality. The reviewed evidence does not establish that the halo effect explains mortality trends. This point needs substantial caution or removal.
- Halo Effect, sentence beginning “This highlights firsthand perception”: Calling perception “the root of the issue” overlooks structural sources of pollution, regulation, housing, income, and access to alternatives. Consider framing perception as one contributing barrier rather than the root cause.
- Halo Effect, sentence beginning “Previous research has examined the effect of the halo effect”: Avoid repeating “effect.”
- Halo Effect, Chile study paragraph: “As cited” is unclear, and the claim about wood-burning stoves across the Global South is very broad. Consider limiting the statement to the location and source being discussed.
- Halo Effect, same paragraph: Persistent pollution despite intervention may suggest several explanations, not only a halo effect. Fuel availability, cost, cultural practice, housing quality, and enforcement may also matter. The interpretation should acknowledge alternative explanations.
- Halo Effect, questionnaire description: Several phrases are incomplete, including “a short five-item evaluating.” Consider summarizing the instrument more concisely and clearly.
- Halo Effect, PM2.5 paragraph: “High levels of particulate matter levels” is repetitive. Units should be formatted consistently as $\mu\text{g}/\text{m}^3$, and the WHO guideline should be checked against the relevant averaging period.
- Page 10, home halo calculation: The explanation of 80%, 70%, and the resulting 10% halo effect is difficult to interpret. Consider clarifying what each percentage represents and whether the difference is statistically significant.
- Page 10, Hofflinger model paragraph: The description provides useful context, but it is lengthy. Consider naming the model type and reporting only the predictors that were significant or central to the interpretation.
- Page 10, neighborhood halo sentence: The sentence is grammatically difficult and the Rankin citation is placed after a period. Clarify the relationship among the home halo, neighborhood halo, and localization of risk.
- Page 10, policy transition beginning “Since the significance”: “Steadily supported” is vague. Consider stating more specifically what the evidence supports and what remains uncertain.
- Page 10, sentence beginning “We could even go as far”: This is conversational and makes a broad claim. Policy effectiveness depends on more than public attitudes alone.
- Halo Effect, real-estate subsidy proposal: This is creative, but the recommendation is not directly tested by the studies reviewed. Consider labeling it clearly as a proposed application and discussing possible

equity concerns, including whether subsidies for fashionable neighborhoods could widen existing environmental inequalities.

- Page 11, Green-Status badge proposal: The idea is understandable, but it needs stronger grounding in evidence on social norms, digital badges, or public recognition. Consider whether privacy, verification, unequal platform access, and performative behavior should be acknowledged.
- Page 11, phrase “people love to gain social capital” and “the in thing”: These expressions are too informal for an academic review and rely on generalization.
- Halo Effect section, overall: The section has a clear structure and a strong environmental application. Its main weakness is that the policy claims sometimes move too quickly from association to intervention. A clearer distinction among evidence, interpretation, and proposal would improve the section considerably.

V. Confirmation Bias

- Page 11, opening paragraph: The concept is introduced accessibly, but the sentence is long and contains a parallel-structure error between “tendency to cherry-pick” and “rejecting.” Consider a more concise academic definition before the quotation.
- Page 11, Nickerson quotation: The comma placement around the author attribution is incorrect. Confirm the page reference and quotation formatting.
- Confirmation Bias, Wason task: The explanation is generally clear, but it could be shortened. The key point is that participants sought confirming cases rather than tests that could falsify their rule.
- Page 11, sentence beginning “Each round”: “Subsequently feedback to whether” is grammatically incorrect.
- Page 11, sentence beginning “The methods participants employed”: “Bore affirmative evidence” is awkward. The interpretation would be clearer if it described how confirmatory testing preserved an incorrect hypothesis.
- Confirmation Bias, Lord study: The study is useful for showing belief polarization. Verify the author names and publication year, since the bibliography lists Lord, Ross, and Lepper (1979), while the text states Lord, Lee, and Lepper (1981).
- Page 12, percentage changes: Explain what the percentages represent and whether they are percentage-point changes or a transformed attitude measure. Without this context, the numbers may be misleading.
- Page 12, sentence beginning “Such data strongly supports”: Subject-verb agreement should be corrected. “Will lead” is also too definite for one study.
- Page 12, paragraph beginning “The confirmation bias is highly relevant”: The logic is not yet clear. The fact that pollution has been recognized for a long time does not itself establish confirmation bias. Consider beginning with the persistence of prior beliefs despite updated evidence.
- Page 12, citation spelling: “Shwartz” and “Schwartz” are both used. Standardize the author’s name and verify the publication year, which also differs between the text and bibliography.
- Page 12, sentence beginning “Previous research, albeit sparse”: “Substantiates the correlation” may be too strong given the observational nature of the cited studies. Consider saying that the research is consistent with or suggestive of confirmation bias.
- Page 12, U.S. air-quality paragraph: The comparison is interesting, but the data appear historical. Consider adding a more recent source or explaining why the older perception data remain relevant to current policy design.
- Page 12, sentence beginning “Ironically enough”: Remove conversational commentary and present the contrast directly.
- Page 12, claim that most American cities enjoy relatively good air quality: This statement is broad and may obscure local disparities, episodic exposure, and pollutants other than ozone. It needs qualification.

- Page 12, sentence beginning “It is likely the presence of confirmation bias”: This is an inference, not a direct finding. Clearly label it as the author’s interpretation and acknowledge other explanations, such as media exposure, uneven local improvement, or distrust of official data.
- Confirmation Bias, King study: The socioeconomic and racial findings are important but sensitive. Take care not to imply that racial composition itself is a simple cause of inaccurate perception. The discussion should distinguish measured neighborhood characteristics, structural inequality, place stigma, and individual bias.
- Page 13, long sentence beginning “Key results are as follows”: The sentence contains several statistics and should be broken into shorter units or presented in a table.
- Page 13, sentence beginning “Thus, social cues”: “Emerge as the key predictors” is too definitive unless the statistical model establishes this after adjustment for objective pollution and other covariates. Consider reporting the study’s actual conclusion more carefully.
- Page 13, “racial/ethnic phenotypes”: This terminology is not appropriate for neighborhood demographic composition. Consider using racial and ethnic composition or another term consistent with the source.
- Page 13, policy transition beginning “Fully grasping”: “Leverage on” should be corrected. More importantly, the ethics of using confirmation bias should be discussed, particularly the difference between value-aligned communication and manipulative targeting.
- Confirmation Bias, conservative and progressive messaging examples: The examples are plausible, but they risk oversimplifying political groups and reinforcing stereotypes. Consider grounding the recommendation in audience-segmentation or framing research and acknowledging variation within groups.
- Page 14, trusted-messenger proposal: This is one of the more practical recommendations. It would be stronger with evidence that trusted messengers improve environmental-health communication or behavior.
- Confirmation Bias section, overall: The section presents strong foundational evidence and a useful policy application. It needs more caution when using observational studies to infer a specific cognitive mechanism, and the ethical limits of targeted messaging should be addressed.

VI. Loss Aversion

- Page 14, opening coin-toss example: The example is memorable, but direct address and rhetorical questions create a conversational tone. Consider summarizing the demonstration in a more formal way.
- Page 14, sentence beginning “From this demonstration”: The claim that people place “twice as much value” on losses should be presented cautiously. The approximate loss-aversion coefficient varies by context and is not a universal constant.
- Page 14, prospect theory paragraph: The theoretical explanation is relevant, but the distinction between loss aversion and risk aversion should be made especially clear. These are related but not interchangeable concepts.
- Page 14, quoted definition: The quotation appears to use “gain” where “gains” may be expected. Verify the wording against the original source.
- Loss Aversion, Karle experiment: The experiment is explained in considerable detail. Consider condensing the lottery mechanics and focusing more directly on what the design shows about risk preference versus loss aversion.
- Page 14, citation year: The text gives Karle et al. (2013), while the bibliography lists 2015. Verify and standardize.
- Page 15, interpretation of Series A: The description of risk aversion appears reversed. Choosing a sure amount below the lottery’s expected value may indicate greater risk aversion, but the classification needs careful checking against the original study.

- Page 15, Series B values: The list says there are six choices but provides seven values of R. Check the original experiment and correct either the number of choices or the list.
- Page 15, sentence beginning “An individual that, additionally”: This is difficult to follow and should be simplified.
- Page 15, environmental-action paragraph: The connection between immediate private costs and future environmental benefits is useful. However, the statement that people choose efficient products only when payback is very short is broad and should reflect the specific scope of the cited review.
- Page 15, sentence stating people value the loss of time more than environmental benefit: This interpretation may conflate time preference, uncertainty, liquidity constraints, habits, and loss aversion. Consider acknowledging competing explanations.
- Loss Aversion, Bartczak study: The study is relevant, but the layout appears broken across lines and needs formatting cleanup. The treatment descriptions can also be condensed.
- Page 16, sentence beginning “The paper provides conclusive evidence”: Avoid “conclusive.” A field experiment can provide evidence consistent with loss aversion, but it does not settle the broader question across all environmental contexts.
- Page 16, Knobloch study description: Calling the study observational is inaccurate if it is a simulation or modeling study. Use the study design described by the authors.
- Page 16, claim that cultural factors do not predominantly influence heating decisions: This needs support and may be difficult to defend across 59 regions.
- Page 16, carbon tax and subsidy paragraph: The explanation of taxes as the “gain element” and subsidies as the “loss element” appears conceptually confused. A carbon tax is generally experienced as a cost, while a subsidy reduces an upfront cost. Recheck the model’s framing and reference points.
- Loss Aversion, tax comparison: The quantitative comparison is useful but should be reported with the model assumptions and scenario context. Avoid generalizing the result beyond the simulation.
- Page 17, phrase “double the tax”: This is too informal and may overstate what the model establishes. State the scenario comparison precisely.
- Page 17, sentence beginning “Having proven”: The section has not proven that loss aversion governs all pollution-related behavior. Use more cautious language.
- Page 17, health-insurance analogy: The cross-domain example is creative, but transfer from health incentives to household emissions should not be assumed. Consider explaining why the mechanism is expected to generalize and what practical differences may limit it.
- Page 17, green-bonus proposal: The recommendation needs discussion of measurement, household size, income, baseline emissions, renter control over appliances, and the risk of penalizing households with limited ability to change consumption.
- Page 17, commuter streak proposal: Capitalization of “Commuters” is unnecessary. The example should also consider accessibility, disability, rural transit availability, and whether losing a fare discount could disproportionately affect lower-income commuters.
- Loss Aversion section, overall: This section has a strong policy focus, but it needs the most technical checking. The manuscript should verify the experimental classifications, distinguish loss aversion from risk aversion and time preference, and reassess how taxes and subsidies are described within the reference-point framework.

V. Discussion

- Page 18, section numbering: This section is labeled V even though earlier sections already use V and VI. Renumber the entire manuscript after the final structure is confirmed.
- Page 18, opening paragraph: The integrative aim is the most original part of the manuscript. The sentence beginning “This section proposes” is grammatically incomplete and should be corrected.

- Page 18, claim of a “perpetual compounding effect”: This is stronger than the evidence supports. The reviewed studies do not directly test the joint operation of all three heuristics. Continue framing the model as a hypothesis or conceptual proposition.
- Page 18, sequential-cycle paragraph: The author appropriately states that the model is not definitive. Consider adding a small conceptual figure showing the proposed sequence and feedback loops, while labeling it clearly as a proposed framework.
- Page 18, halo-effect paragraph: The theoretical rationale for placing the halo effect first is understandable, but the sequence may differ by context. Prior beliefs or loss concerns could precede place attachment. Discuss alternative ordering rather than implying one universal pathway.
- Page 18, phrase “in a more favorable direction that objective data warrants”: “That” should be “than.”
- Page 18, sentence beginning “The critical point”: This is a comma splice and should be divided or joined correctly.
- Page 18, confirmation-bias paragraph: The examples are clear and help the reader understand the proposed mechanism. Because they are hypothetical examples, make that status explicit.
- Discussion, loss-aversion paragraph: The transition is logical, but “functions are the final barrier” contains a grammatical error. “Corrected perception” and “corrected behavior” also sound overly binary; behavior can change partially.
- Page 19, synthesis paragraph: “More intractable that” should be “more intractable than.” Consider stating what empirical design would be needed to test whether the combined effect is additive, interactive, or sequential.
- Page 19, phrase “suggestive, if not conclusive”: Since the manuscript later acknowledges that the cases do not prove interaction, “if not conclusive” should be removed.
- Page 19, Chile case: The synthesis is thoughtful, but the interpretation attributes persistence to confirmation bias and loss aversion without direct measures of either mechanism. Label these as plausible explanations and retain alternative structural explanations.
- Page 19, “Padre Las Casa”: Check the place name. The earlier section uses Padre Las Casas.
- Page 19, “systemically overestimated”: The intended word appears to be “systematically.”
- Discussion, Chicago case: “Aur quality” is a typographical error. More importantly, the claim that loss aversion discouraged investment is speculative and is not measured in the cited study. This should be presented as a research question rather than an inferred mechanism.
- Page 20, policy implications paragraph: The example of belief change without behavior change is useful. Consider linking this distinction to established models such as the intention-behavior gap if relevant literature supports it.
- Page 20, final paragraph: The last sentence repeats “potentially complementary component of a broader behavioral strategy” and appears unfinished. The proposed sequence also stops after confirmation bias and does not complete the loss-aversion component.
- Discussion, overall: This section provides the manuscript’s clearest original contribution. It should be expanded as a conceptual framework but separated carefully from the empirical findings. A short subsection on testable propositions, alternative pathways, and research design would make the contribution much more convincing.

VI. Conclusion

- Page 20, opening sentence: “Compelling evidence” and “profoundly shape” are stronger than warranted across all three domains. Consider stating that the reviewed evidence indicates or suggests meaningful influence.
- Page 20, punctuation around heuristic list: The manuscript uses dash characters inconsistently. Replace them with commas or parentheses and standardize punctuation throughout.

- Page 20, sentence beginning “Findings explored”: The wording is slightly rhetorical. Consider summarizing the practical implication directly.
- Conclusion, heuristic-summary paragraph: This is a useful recap, but several claims remain causal. For example, confirmation bias may contribute to persistent perceptions, but the observational studies do not isolate it experimentally.
- Page 21, phrase “less than proportionate social impacts”: Clarify what kind of impact is meant, such as public recognition, trust, behavior, or policy support.
- Page 21, interaction paragraph: “Prevents updates perception” should be corrected. “Casual orderings” should be “causal orderings.”
- Page 21, future-research sentence: This is a strong and necessary limitation. Consider moving part of this caution earlier into the Discussion so the speculative model remains appropriately framed throughout.
- Page 21, policy paragraph: “Effective countervailing policies work with, not against” reads as a settled conclusion. Consider saying that policies may be more effective when they account for these patterns.
- Page 21, sentence beginning “As proposed, this can involve”: The transition is awkward. The three proposals could be presented as a concise list or parallel sentences.
- Page 21, loss-framed incentives: The example of forfeited bonds raises ethical and equity questions. Consider acknowledging that loss-framed designs may burden vulnerable groups or reduce trust if they are perceived as punitive.
- Conclusion, final paragraph: “Beget real, effective action” is overly rhetorical. The conclusion would be stronger if it ended with a precise statement of what this review contributes and what evidence is still needed.
- Conclusion, overall: The conclusion accurately reflects the paper’s main themes, but it should distinguish more carefully among established evidence, conceptual synthesis, and policy proposals. This distinction is essential because the manuscript’s strongest contribution is the proposed integration, not proof that the sequence already operates as described.

Figures, Formatting, Citations, and Editorial Consistency

- Figure 1 in Halo Effect section: The figure is relevant, but the image quality is low. Replace it with a clearer version if permission allows, or recreate the figure based on the reported data with proper attribution.
- Figure 1 caption: Use a consistent citation format and change “Pg” to the journal’s preferred page notation.
- Section numbering: The manuscript has duplicate and incorrect Roman numerals. Renumber sections after all structural revisions are complete and ensure the roadmap matches the final numbering.
- Citation style: The manuscript mixes author-date citations with numbered bibliography markers and stray superscript numbers. Choose one style and apply it consistently.
- Spelling convention: British and American spellings are mixed, including behavioural/behavioral, organisation/organization, judgement/judgment, neighbourhood/neighborhood, and maximise/maximize. Select one convention based on journal requirements.
- Punctuation: Several dash types appear throughout the paper, and some have encoding problems. Replace them with standard punctuation and review spacing around all dashes, parentheses, and citations.
- Numbers and units: Standardize percentages, currency, PM2.5 notation, $\mu\text{g}/\text{m}^3$, euro symbols, and CO2 notation. Ensure spaces and superscripts follow one style.
- Quotations: Verify all direct quotations against the original sources and include page numbers consistently.

- Source accuracy: Several years and author names differ between the manuscript and bibliography. Cross-check every in-text citation against the reference list.
- Bibliography: Although the references were not reviewed line by line here, the list should be standardized for capitalization, author formatting, dates, URLs, DOI presentation, and hanging indentation before publication.

Overall Feedback

- This manuscript addresses a timely and important problem by bringing behavioral economics into the study of air pollution perception and policy. The interdisciplinary focus is a real strength. The author has also selected three heuristics that are understandable, policy-relevant, and distinct enough to support a clear structure. The use of foundational psychological studies followed by environmental applications helps readers see how each concept developed and why it may matter outside the laboratory.
- The strongest contribution is the Discussion section, where the author proposes that the halo effect, confirmation bias, and loss aversion may reinforce one another. That idea gives the paper more value than a standard summary of three separate biases. At present, however, the interaction model remains conceptual. The case studies reviewed do not directly measure all three heuristics at once, so the manuscript must consistently distinguish between what the evidence shows and what the author is proposing. Developing the model as a set of testable propositions would strengthen the paper and create a clearer direction for future research.
- The manuscript also needs more careful handling of causal language. Several observational studies are used to attribute perceptions or behavior to a particular heuristic, even though other explanations are possible. Structural conditions, socioeconomic constraints, housing quality, access to alternatives, political trust, and media exposure may shape the same outcomes. A stronger review would acknowledge these competing explanations rather than treating cognitive bias as the sole or primary cause.
- The policy sections are creative and practical, but some recommendations move beyond the evidence reviewed. The author should label original proposals clearly and discuss feasibility, ethics, and equity. Incentives based on loss aversion, neighborhood status, or targeted political messaging may affect groups differently and may create unintended consequences. Adding this critical dimension would make the recommendations more credible and responsible.
- In terms of writing, the manuscript has a clear voice and shows genuine engagement with the topic. The next revision should focus on reducing repetition, shortening detailed summaries of classic experiments, correcting grammar and citation inconsistencies, and maintaining a formal academic tone. More space should be devoted to comparing findings, evaluating limitations, and explaining where the evidence is strong or uncertain.
- Overall, the paper has a promising argument and a potentially original conceptual contribution. Substantial revision is needed before publication, particularly in the areas of methodological transparency, scientific caution, policy justification, section organization, and editorial consistency. With careful revision, the manuscript could become a thoughtful and useful review of how behavioral processes influence environmental risk perception and policy response.

Final Recommendation

- **Revise and submit.** This manuscript addresses a timely and important topic, and its interdisciplinary approach brings together behavioral economics, environmental risk perception, and policy design in a thoughtful way. The discussion of the halo effect, confirmation bias, and loss aversion provides a strong foundation for exploring how cognitive biases influence environmental decision-making. The proposed conceptual framework is an interesting contribution and gives the review a clear identity beyond simply summarizing the existing literature.

- The comments above are intended to strengthen an already promising manuscript. Most of the suggested revisions focus on improving clarity, strengthening the connection between the evidence and the conclusions, increasing methodological transparency, and distinguishing more clearly between findings supported by the literature and the author's original interpretations and policy proposals. Addressing these areas will make the central argument even more compelling and easier for readers to follow.
- From an editorial standpoint, I also recommend maintaining one spelling convention throughout the manuscript. Please choose either American or British English and apply it consistently across the text, headings, figure captions, tables, and references. Consistency in style contributes to the overall professionalism and readability of the paper.
- The paper already has many strengths, and I hope the suggestions provided are helpful as you continue refining it. I appreciate the amount of work that has gone into this review. Bringing together literature from multiple disciplines is challenging, and the manuscript reflects careful research and genuine engagement with the topic. With the revisions outlined above carefully considered and incorporated, I believe the manuscript will be substantially strengthened and suitable for resubmission.

Review of: The Psychology of Pollution: Analyzing How Heuristics Affect Perception of Air Quality and Policy Design

Recommendation: Accept with revisions

Overall assessment

This manuscript addresses an important question of how cognitive biases shape public understanding of air pollution and the effectiveness of environmental policies. The author draws from various sources including behavioral economics, psychology, environmental health, and public policy. The discussion on the interacting effects of three heuristics is creative and is the paper's most original contribution.

The manuscript is promising and demonstrates strong engagement with the relevant literature, but some revisions are needed before publication. The primary areas to improve include: distinguishing between evidence of perceptual gaps vs evidence of specific cognitive biases, clarifying relationship between perception and behavior across the 3 heuristic sections, further development and more careful framing of the proposed model as a conceptual in the discussion, and considering limitations and ethical implications of the proposed policy interventions across all sections.

Major comments:

- In the section on the halo event, the proposed policy implications are creative but the mechanism requires clarification. The intervention relies less on the traditional halo effect (i.e., existing positive impressions influence judgments of unrelated attributes) and more on social signaling/identity-based motivation, where sustainability becomes associated with status and desirability. A more appropriate example of the halo effect would be for example that affluent neighborhoods are assumed to have cleaner air/better environmental quality even when it is not the case. I suggest that the author considers this distinction and alters the proposal to be more in line with the halo effect or provide evidence supporting the assumption that exposure to green infrastructure in prestigious places leads to broader behavioral adoption. Additionally, there are ethical/equity concerns to consider here. Investments in green infrastructure in affluent neighborhoods would disproportionately benefit already advantaged communities – and in fact this is already happening today. Consider reframing the intervention to ensure these behavioral insights are used to increase adoption of sustainable infrastructure across different communities, especially those experiencing the greatest pollution burdens. At minimum, consider mentioning this nuance.
- For the section on confirmation bias, the author gives a strong overview and does a good job connecting foundational psych work to perceptions of the environment. However, be careful in distinguishing between evidence of misperception vs confirmation bias. E.g., a discrepancy between objective measures and subjective perceptions of air quality doesn't necessarily show confirmation bias, as there are many alternative explanations such as differences in personal experience, distrust in monitoring systems/science, media exposure, changing expectations/standards, or attention. To more directly support the confirmation bias argument, consider focusing more on evidence that shows people

selectively seek out, interpret, or remember information in ways that reinforce pre-existing beliefs. Additionally, I'd stay away from sing phrases like "leveraging confirmation bias" as it is typically thought to be a negative barrier that stops people from processing info accurately. You don't want to seem as if you suggest exploiting confirmation bias but rather that policymakers can work with existing values/identities to improve receptiveness to information. Relatedly, the author may want to acknowledge possible risks of targeting messaging based on community as it may unintentionally deepen divisions by allowing different groups to receive substantially different narratives about the same issue. Consider discussing how identity-based communication can be used ethically while still maintaining scientific consistency.

- The loss aversion section makes an interesting argument about how loss aversion affects environmental behavior, but its connection to the paper's broader focus of air pollution perception is less clear. Unlike halo effect and confirmation bias sections, this section primarily addresses behavioral responses rather than how people perceive pollution risks. Consider explicitly explaining whether loss aversion acts as a barrier after perception has already changed (i.e., people recognize pollution as a problem but avoid costly solutions), rather than presenting it as another mechanism shaping perception itself. Also, the proposed applications to environmental policy are interesting, but should discuss possible limitations of these approaches. For example, penalties or forfeitable incentives may disproportionately affect households with fewer resources or limited ability to adopt greener technologies
- I think this proposed interaction model is the most original aspect of the manuscript, but requires more careful framing. The studies reviewed earlier provide evidence that each heuristic may influence perceptions/behavior independently, but don't show these biases operate sequentially or causally in the order proposed by the author. For example, the halo effect does not necessarily need to precede confirmation bias, and loss aversion may influence whether individuals seek or interpret information in the first place. I suggest softening the framing of this section to frame it as a conceptual framework or hypothesis rather than a proposed sequence supported by existing evidence. Also, you might consider strengthening this section by outlining specific predictions that future empirical work could test. Alternatively, consider if these are better conceptualized as interacting factors rather than fixed sequence. A cyclical model may better capture that they may reinforce each other in multiple directions instead of occurring in order.
- The conclusion is generally strong and appropriately synthesizes the paper, but overreaches on some claims (similarly to earlier sections). I suggest softening some of the claims to better reflect the strength of the evidence presented by using language such as "may substantially influence" or "can contribute to" rather than definitive, established mechanisms.

Minor comments:

- "Economic costs amounting to trillions of dollars—eroding national GDPs and straining healthcare systems—are starkly documented." □ no citation

- “These deviations from neoclassical economics are instances of what is known as behavioural economics.” □ Improve transition from previous paragraph, defining neoclassical economics should come before saying “these deviations from neoclassical economics” – or should be incorporated into that first sentence
- I would suggest incorporating more evidence or examples into first paragraph of intro
- Introduction is under-cited
- Unclear: “There exists a growing body of literature on both air pollution and behavioural economics individually, but a critical gap persists at their intersection. Where behavioural approaches have been recently applied to environmental risk perception, they tend to examine cognitive biases within a narrow empirical scope, rather than developing an integrated account from various contexts of how the general bias impacts air quality perception as a whole. The systemic application of heuristics—namely the halo effect confirmation bias, and loss aversion—to air quality perception and pollution-related policy design therefore remains underdeveloped in existing literature.”
- The last 6 paragraphs of the intro which outlines the structure of the paper is good but long and slightly repetitive so it should be condensed
- Consider including “narrative review” in title or abstract so it is clear earlier on the structure that it will take and why there is a methods section
- Not necessary to say “Studies cited later on in this paper will focus on levels of PM_{2.5} as a measure of air quality due to the seriousness of its health consequences and geographical ubiquity (Henning, 2024).” – this is talked about in paragraph after and can speak for itself
- Font and size changes throughout, particularly in Loss Aversion section
- Citations are often placed outside punctuation (i.e., . (Rankin, 1969))
- “I suggest that governments can provide subsidies or tax rebates to real estate projects” – first use of first person, either introduce throughout or change
- Informal wording here:
 - “**We could even go as far to say** that the success of mitigation policies hinge on thorough understanding...”
 - “**Since people love to gain social capital and follow social trends**, being eco-conscious will come to be seen as the “**in**” **thing**, encouraging widespread adoption of green behaviours through social proof.”
- Be cautious of overstating causality:
 - “The confirmation bias is highly relevant in the context of air pollution, **lying in the fact** that air pollution has been a recognised problem for millennia”
- “Up till this point” – should be until
- “tap on this heuristic” should be “tap into this heuristic”
- Wording is odd here: “These examples serve to nudge individuals to more willingly adopt greener lifestyle habits.”
- “The paper provides conclusive evidence for loss aversion in the environmental context.” – too strong, these findings provide evidence consistent with loss aversion influencing environmental decisions, but other psychological and contextual factors may also contribute

The Psychology of Pollution: Analysing How Heuristics Affect Perception of Air Quality and Policy Design

Acknowledgements

Keywords:

Pollution, Air quality, Behavioural economics, Confirmation bias, Halo effect, Loss aversion, Policy ~~d~~Design, Risk perception

Abstract

Air pollution is a severe, global issue responsible for millions of premature deaths yearly and economic costs totalling trillions of dollars. Heuristics—mental shortcuts individuals intuitively employ to aid decision-making—are found to affect public perception of this issue due ~~to its the~~ subjectivity and complexity ~~of individual discernment based on visual and olfactory cues~~. This paper ~~conducts an analysis into~~examines the critical role of three key heuristics (the halo effect, confirmation bias, and loss aversion) in shaping public perception, often leading to a significant gap between objective reality and subjective experience. The halo effect ~~causes can be linked to~~ residents ~~to~~ systematically overestimate~~ing~~ the cleanliness of air in their own homes due to emotional place attachment shaping global evaluations of their living environments. Confirmation bias explains the persistence of entrenched negative ~~(or positive)~~ perceptions of air pollution in the face of objective environmental improvements ~~(or decline)~~. Loss aversion hinders the adoption of energy-efficient technologies ~~as upfront costs are weighted more heavily as losses than future energy savings are weighted as gains and reduces the effectiveness of policies that trigger individuals' loss aversion~~disproportionately undermines policies that ~~do not directly offset upfront investment costs, such as carbon taxes~~. Through a ~~narrative~~ review of foundational psychological experiments and environmental studies, this analysis demonstrates how these heuristics systematically distort risk assessment and hinder effective public response to pollution. The paper concludes by proposing targeted urban, communication, and incentive policies designed to either ~~mitigate and/or strategically harness these biases~~.reduce the effects of these biases or harness them to improve policy response.

The Psychology of Pollution: Analysing How Heuristics Affect Perception of Air Quality and Policy Design

I. Introduction

Air pollution stands as one of the most pervasive and severe environmental ~~causes-threats~~ of the 21st century, ~~imposing a staggering toll on global public health and economic stability (WHO, 2016)~~. Millions of premature deaths annually are attributed to exposure (WHO, 2016), ~~taking a toll on global public health~~. Economic costs amounting to trillions of dollars—eroding national GDPs and straining healthcare systems—are starkly ~~documentations of pollution-induced economic instability~~. Yet, societal and individual responses to this ~~singular~~ threat are not dictated solely by these physical and epidemiological facts. Instead, they are filtered through the complex lens of human perception, a process fraught with cognitive shortcuts (i.e. heuristics) that can systematically distort our understanding of risk.

~~These deviations from neoclassical economics are instances of what is known as behavioural economics. Generally, economic theory rests on the n~~Neoclassical economics ~~assumption thates most rational actors have stable preferences and make well-informed decisions with perfect information maximisein self-interest personal utility, while~~ ~~b~~Behavioural economics ~~deviates from this assumption by~~ incorporat~~ing~~es psychology, and account~~ing~~s for real-world inconsistencies in rationality caused by cognitive biases, emotions, and context. Behavioural economics explains that people are often unable to fully analyse situations involving economic and probabilistic judgements (~~Kahneman and Smith, 2002~~). Individuals thus tend to rely on heuristics to form judgements, which can be systematically biased. (~~Kahneman and Smith, 2002~~)

This paper argues that ~~to effectively address the air pollution crisis, we must movinge~~ beyond neoclassical economic analysis and engage~~ing~~ with the behaviou~~ral~~ economics of risk perception~~—~~, where heuristics play a decisive and often counterproductive role~~—can~~ ~~raise the effectiveness of anti-pollution movements.~~

The imperative for this behaviou~~ral~~ approach stems from the fundamental nature of the threat itself. ~~The data and sources cited in this paper focus on the PM_{2.5} pollutant (i.e. fine particulate matter), which is~~The most harmful pollutants, such as fine particulate matter (PM_{2.5}), are imperceptible to the human senses. Unable to directly “see” the danger, individuals are forced to construct judgements based on imperfect proxies—visual haze, olfactory cues, ~~media reports~~, or physical discomfort. ~~Socially supplied information, such~~

Formatted: Not Superscript/ Subscript

as that from media reports and hearsay, can distort judgements as well. This complex cognitive task, laden with uncertainty and complexity, triggers our reliance on our tendency to rely on intuitive mental shortcuts to simplify tasks. These heuristics are essential for efficient decision-making but introduce predictable biases. In the context of an invisible, diffuse, and scientifically complex hazard like air pollution, the gap between objective data and subjective perception widens. This, making public comprehension and systemic prioritization/prioritisation of the issue vulnerable to systematic error.

Drawing this connection between environmental perception and heuristic-driven distortion has critical implications for understanding these cognitive processes is not a simple academic exercise but rather a critical component of policy design. If public concern, personal protective behaviour, and political will are shaped more by heuristic-driven perceptions than by instrumental data/objective measurements, then policies based solely on perfect information or rational economic incentives may fail to resonate or motivate action. Behavioural economics provides the framework to decode this disconnect. By examining how specific biases operate within the domain of environmental risk, we can may be able to better anticipate public response, design more impactful/effective communication strategies, and craft interventions that align/correspond to how people actually process information and make choices.

There exists a growing body/collection of literature on both air pollution and behavioural economics individually, but a critical/key gap persists at their intersection. Where behavioural approaches have been recently applied to environmental risk perception; however, they tend to examine cognitive biases within a narrow empirical scope, rather than developing an integrated account from across various/multiple contexts of how the general bias impacts air quality perception as a whole (He et al., 2018; Liu et al., 2018) (He et al., 2018; Liu et al., 2018). The systemic application of heuristics—namely the halo effect, confirmation bias, and loss aversion—to air quality perception and pollution-related policy design therefore remains underdeveloped in existing literature. This paper responds to that gap by examining each of these three heuristics in turn, synthesising psychological and environmental evidence to demonstrate how each independently distorts public understanding of air quality and undermines conventional policy responses. z

This paper first provides background on the physical characteristics of air pollution and its documented health and economic impacts, emphasising the challenge of perceiving an invisible risk. It then analyses three key heuristics—the halo effect, confirmation bias, and loss aversion—in the context of air pollution. This paper proceeds by first establishing the necessary background on air pollution's physical characteristics and its documented health and economic impacts, while highlighting the inherent challenge of perceiving an invisible

risk. It then delves into an analysis of three key heuristics—the halo effect, confirmation bias, and loss aversion—and their relevance to air pollution.

The background section outlines the objective reality of air pollution, citing recent global and national data that call for urgency, and detailing its detrimental health consequences. It also establishes why the perception of air pollution is a psychological puzzle, necessitating a behavioral inquiry.

The analysis on the halo effect explores how an overall positive (or negative) impression of a place, such as one's home or neighborhood, creates a cognitive bias that leads individuals to systematically overestimate (or underestimate) local air quality, even when faced with contrary objective data. Drawing on foundational behavioral economics studies and contemporary environmental research, such as investigations into wood-burning stove usage in Chile (Boso et al., 2020), this section demonstrates that overall neighbourhood satisfaction creates a “home halo”, leading residents to perceive indoor and local air as cleaner or dirtier than measurements indicate.

An examination of the confirmation bias investigates how individuals selectively interpret information about air pollution to reinforce pre-existing beliefs, leading to a persistent gap between improving objective data and stagnant or polarized public perception. The section dissects past studies on U.S. air quality trends and neighbourhood stigmas in Chicago to illustrate how deep-seated beliefs and social stigmas cause people to overlook empirical evidence of improvement or risk (King, 2014).

The discussion of loss aversion applies the principle of the disproportionate weight given to potential losses—compared to equivalent gains—to the slow adoption of pro-environmental behavior and technologies. This section draws on experimental evidence and applied models, such as those analyzing wildfire risk perceptions and residential heating choices (Bartczak et al., 2015), to demonstrate that individuals' risk preferences become markedly more conservative when decisions involve perceptible losses, such as high upfront costs or lifestyle changes.

The paper proceeds as follows. In Section III, we clarify the methodology for study selection and data extraction. In Section II, we delve deeper into take a closer examination of the facts behind air pollution. In Sections IVH, IVH, and IVI, we explore the halo effect, confirmation bias, and loss aversion, respectively. Section VII concludes.

II. Methodology

This paper adopts a narrative review framework to synthesise existing empirical and theoretical literature on the role of cognitive heuristics in shaping public perception of air quality. A narrative approach was selected on the grounds that the primary objective of this review is thematic synthesis and theoretical integration across disciplines, namely behavioural economics, environmental science, and public policy, rather than an exhaustive enumeration of all available studies, and is thus interpretive instead of systematic.

Literature was identified through searches of Google Scholar, PubMed, Web of Science, and JSTOR, supplemented by reports from international organisations including the World Health Organisation (WHO), the United Nations Environment Programme (UNEP), and the Institute for Health Metrics and Evaluation (IHME), conducted from June 2025 to November 2025. Search terms included combinations of the following: “air pollution”, “halo effect”, “confirmation bias”, “loss aversion”, “air quality perception”, and “cognitive bias”, and “environmental policy”. The same search combinations were used across databases. No strict date restriction was imposed, as the review deliberately includes foundational experimental studies and theoretical papers, such as Kahneman and Tversky’s paper on prospect theory (1979) and Wason’s paper on the rule discovery task (1960), providing a core, well-tested psychological base to build upon. More recent empirical studies, prioritised from 2000 onwards, were selected to ground these psychological frameworks in contemporary environmental contexts.

Studies were included based on relevance to at least one of the three heuristics under examination, methodological transparency and publication in peer-reviewed journals or by credible globally recognised international institutions; purely anecdotal or non-empirical commentary was excluded from citation. Findings were then organised thematically; for each heuristic, foundational experimental evidence establishing the psychological mechanism was identified first, followed by domain-specific studies applying that mechanism to air pollution perception or responsive behaviour. Policy implications were derived inductively from the patterns observed across studies within each thematic cluster, with the aim of translating behavioural insights into actionable regulatory recommendations. Simple policy recommendations were then constructed to address the identified gaps directly.

III. Background

As touched upon in the introduction, air pollution represents one of the most significant environmental health risks of the modern era, with profound implications for public health and economic productivity. Simultaneously, perception of this risk is a complex

psychological process, and may be filtered through cognitive biases, transforming measured pollution levels into varying before becoming a personalised human interpretation of objective reality interpretations from individual to individual. This section outlines the fundamental nature of air pollution as a physical phenomenon, its proven impacts, and the nuance involved in human perception, thereby establishing the necessity of investigating the heuristics that shape this perception.

Air pollution is the “contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere” (WHO, 2024, p. 1). Studies cited later on in this paper will focus on levels of PM_{2.5} as a measure of air quality due to the seriousness of its health consequences and geographical ubiquity (Henning, 2024). However, other key pollutants consist of particulate matter (PM_{2.5} and PM₁₀), which can penetrate deep into the lungs and bloodstream; nitrogen dioxide (NO₂) and sulfur dioxide (SO₂), primarily from combustion processes; and ozone (O₃), a secondary pollutant formed in the atmosphere. Perception of air pollution varies from pollutant to pollutant given their different physical characteristics. Studies cited later on in this paper will focus on levels of PM_{2.5} as a measure of air quality due to the seriousness of its health consequences and geographical ubiquity (Henning, 2024).

The World Health Organisation (WHO) has declared air pollution a public health emergency, with millions of premature deaths annually attributed to exposure (WHO, 2016). As a whole, air pollution accounts for 1 in 8 deaths around the world at the time of writing (2025), and the culprit of 7.9 million deaths in 2023 alone. Regionally, More than 90% of East and Southeast Asia’s population currently breathe air with unsafe levels of PM_{2.5} and ground-level ozone (UNEP, 2024). Additionally, disease outcomes are severe: aAccording to the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) 2023, particulate matter pollution on its own was the second leading risk factor for early death and poor health globally (IHME, 2023). In particular, studies have shown PM_{2.5} exposure is associated with to the increased risk of several diseases such as cardiovascular and respiratory diseases, lung cancer, stroke, pneumonia, depression, and diabetes (Sukuman et al., 2023). Exposure has also been linked to several other conditions, such as asthma, bronchitis, breast cancer, etc., but concrete causal linkage warrants further investigation lung cancer, stroke, pneumonia, depression, and diabetes from short- and long-term exposure (Sukuman et al., 2023).

Apart from destructive health effects, air pollution also weighs heavily on countries’ balance sheets. Globally, Estimates project a 2.9 trillion USD global economic cost, equivalent to 3.3% of the world’s GDP. PM_{2.5} was found single-handedly responsible for 1.8 billion days of work absence, 4 million new cases of child asthma and 2 million preterm births in 2018 (Myllyvirta, 2020). In the UK, air pollution links to around 30,000 deaths per

year are linked to air pollution, along with the depletion of £27 billion due to healthcare costs, productivity losses, and reduced quality of life, or closer to £50 billion if trickle-down effects like dementia are accounted for (RCP, 2025). The UK stands to benefit by £1.6 billion annually if WHO guidelines for air pollution are met (CBI, 2020). In India, air pollution cost an estimated \$95 billion (3% of the country's GDP) in the form of reduced productivity, work absences (3.8 billion working days), and premature deaths (responsible for 18% of deaths in India) in 2019. Had the country managed to half pollution in the last quarter of a decade, GDP is believed to will have risen by 4.5% (Dalberg, 2021). In Southeast Asia, deaths attributed to air pollution are projected to rise 10% by 2050, with costs snowballing to a hefty amounting to 600 billion USD. This figure captures indirect costs such as lost productivity, economic output, and premature deaths. Undoubtedly, good air quality is a key pillar of a strong economy.

While the physical, and sociological facts health and economic evidence of air pollution are critical, they do not, in themselves, dictate public response. The perception of air quality is inherently challenging because individual particles of the most harmful pollutants, such as PM_{2.5}, are invisible to the naked eye and renders ambient haze an imperfect proxy which was termed in reference to its size of 2.5µm or smaller, are often invisible to the naked eye. Moreover, pollution-related health consequences often trigger delayed physical symptoms. In the absence of perfect sensory information, individuals cannot directly “see” the threat and must instead construct a judgement of an invisible environmental factor—air quality (Fang et al., 2025).

To navigate a world full of such complex and uncertain judgements, the human brain relies on a powerful set of mental shortcuts known as heuristics or cognitive biases. Kahneman (2011) uses a conceptual framework of two “systems”—System 1 and System 2—in the human brain to explain this.¹⁵ This operation is a demonstration of our intuitive “System 1” in action, one half of two distinct cognitive systems involved in decision-making as introduced by Kahneman. The system is fast, automatic, and effortless, allowing us to make countless decisions without the laborious mental effort of our analytical “System 2”. Heuristics are the rules of thumb System 1 generally employs—they are generally typically effective but come with a trade-off: they can systematically lead us away from purely rational assessment and towards predictable errors.

These cognitive processes are not a sign of intellectual failure but rather a feature of our mental machinery evolved for efficiency. When faced with the daunting task of assessing an invisible danger like air pollution, it is cognitively more efficient to substitute the hard question “What is the current concentration of PM_{2.5}?” with easier ones like “Does the air look hazy?” or “Does it smell bad?”. Common proxies for pollution levels include visual cues (haze or smog), olfactory cues (smell of smoke, exhaust, or industrial emissions), and

Formatted: Not Superscript/ Subscript

contextual and somatic cues (media reports, seeing people wear masks, or experiencing physical symptoms like coughing or eye irritation). However, these proxies are imperfect and can be misleading—a clear, sunny day may still have dangerously high levels of ozone, while a foggy day might be mistaken for severe smog. This reliance on physical cues makes the perception of air quality highly susceptible to an unconscious use of cognitive biases as it facilitates a shift from a data-rich to cue-driven decision-making process and introduces perceptual vulnerabilities (Boso et al., 2025).

Consequently, the public's perception of air quality is not a direct reflection of objective data but rather a heuristic-driven judgement, which in turn shapes personal behavior, public concern, and political will. In the following three sections, I further probe into three specific cognitive biases: the halo effect, confirmation bias, and loss aversion.

IV. Halo Effect

If you were to meet another individual for the first time and they were treated warmly and friendly by them, you would likely associate them with several other positive qualities, without any real evidence to support those assumptions. This is due to the human tendency to base our general judgement of a person or object off from a single lasting impression of them. We can apply this in a general context where observation of a single positive or negative trait impels judgement in every other aspect. Formally, this is known as the 'halo effect'.

The halo effect was first coined and described by American psychologist Edward Thorndike in his personality studies and has since been one of the most widely known and discussed cognitive biases. Nisbett and Wilson (1977, p. 1) outline the phenomenon as 'the influence of a global evaluation on evaluations of individual attributes of a person (or object)'.⁴⁶ We can understand the heuristic as a mental shortcut that causes an alteration in judgement when a positive or negative impression of a person, place, or object in one area influences our overall perception and assessment of them, even in unrelated areas.

Its existence has been tested and studied at length. Notable evidence comes from Nisbett and Wilson's experiment (1977) on college students, to whom they played two different who watched videotaped interviews of the same instructor. Participants were assigned one of two treatments—one where the instructor was warm and friendly, one where the instructor was cold and distant—then, after viewing, the students were asked to rate the instructor's likeability, as well as physical attractiveness, mannerisms and accent. The halo effect was measured quantitatively by calculating the difference in average ratings for each category between warm and cold treatments.

The between-subject study yielded more favourable ratings in all categories for the warm treatment and consistently poorer ratings for the cold treatment as displayed in the bar chart below.

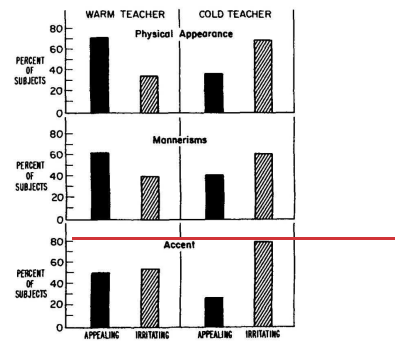


Figure 1: Ratings of Teacher's Physical Appearance, Mannerisms, and Accent as a Function of Manipulated Warmth vs Cold (Nisbett et al., 1977, Pg 253)

As the top third of Figure 1 shows, a sizable majority of warm treatment subjects rated the instructor's physical appearance as appealing, while a sizable majority of the cold treatment subjects rate the instructor's physical appearance as irritating. The central part of Figure 1 shows that ratings of the instructor's mannerisms are similar. Lastly, the bottom third of Figure 1 shows that about half of warm treatment subjects rated the instructor's accent as appealing, while an overwhelming majority of cold treatment subjects rated the instructor's accent as irritating. The warm treatment produced consistently higher ratings across all measures. In particular, warm-treatment participants judged the instructor's appearance and accent more favourably. The results quantitatively prove supporting the existence of a halo effect in play by showing a measurable shift in attribute judgment driven by an unconscious, global evaluation, even when objective information should allow for independent assessment the evaluation of traits unrelated to the evidence presented. However, this person-perception experiment does not draw a direct link to place-based air-quality judgements.

Fundamentally, the halo effect causes our general impression of an entity to shape our view of its unknown attributes. This is especially relevant. The relevance of this heuristic to in air quality perception is due to the ambiguity of air quality as an attribute. If one is emotionally attached to the place where they live (quantified by the difference in perception of perceived air quality between home and city), the positive feelings evoked

when thinking about their homes would subsequently colour perceptions of air quality as clean and favourable (Hofflinger et al., 2019). Since global evaluations ~~are~~ can be powerful enough to ~~trump~~ override other available objective data ~~available~~, the phenomenon likely explains why people in highly polluted cities still perceive their neighbourhood air quality positively. As a result, vigorous public awareness campaigns and a wide array of comprehensive research into the ~~negative medical repercussions~~ adverse health effects of poor air quality have yet to ~~make a dent in prevailing high mortality rates due to air pollution~~ inspire a systemic overhaul of pollutive practices. ~~This is as~~ This highlights firsthand perception of surrounding air quality as the root of the issue. ~~If~~ if people do not perceive the air they inhale as hazardous, they will not make an effort to rectify their habits to protect themselves (D'Antoni et al., 2017).

Previous research has examined the ~~effect~~ impact of the halo effect on individual perception of air pollution in specific contexts.

One such observational study was conducted in Southern Chile (Boso et al., 2020). As cited; wood-burning stoves are one of the main culprits of the high rates of pollution in ~~countries across the Global South~~ Chile. Despite the presence of a wide range of technological options to replace stoves ~~in Chile~~, it still exhibits similarly high levels of pollution. Outdoor and indoor pollution persists despite aggressive government intervention including measures such as refurbishment subsidies and prohibitions of firewood usage, suggesting the involvement of a halo effect. Although, alternative explanations such as fuel availability cost, culture practice, and enforcement, are to be acknowledged as well. The study zeroes in on indoor air ~~pollution to touch on reports, noting~~ that people tended to spend more than 80% of their time indoors.

Residents across 81 homes in Temuco and Padre Las Casas filled in questionnaires on perceived indoor and outdoor air quality ~~on their individual perception of air quality at home~~, which was then compared to on-site sensors and evaluations made by inspectors. ~~The questionnaires consisted of 3 sections. First, a short five-item evaluating indoor and outdoor (neighbourhood and city) at that moment on a nine-point Likert scale. Then, an extended 25-item on air quality in general and on that day. Additional questions on severity of air pollution (five-point Likert scale), their level of concern for its effects on their health (five-point Likert scale), sociodemographic questions and home characteristics finished up the questionnaire.~~ The questionnaire included brief ratings on current air quality, a longer survey on general air quality, and additional questions on pollution severity, health concern, demographics, and home characteristics.

Air healthiness was defined by the threshold of 25 ~~The general environment of the study has high levels of particulate matter levels (PM-level hereafter)—the recommended WHO level~~

is $25\mu\text{g}/\text{m}^3$ based on recommended WHO levels (Unhealthy if $>25\mu\text{g}/\text{m}^3$ and Healthy if $\leq 25\mu\text{g}/\text{m}^3$), while the detected indoor and outdoor $\text{PM}_{2.5}$ levels were reached peaks of $144\mu\text{g}/\text{m}^3$ and $74\mu\text{g}/\text{m}^3$ respectively, and had daily means of $33.07\mu\text{g}/\text{m}^3$ and $35.37\mu\text{g}/\text{m}^3$ respectively, both falling under Unhealthy by objective measurement. However, on a scale of 1–9, average evaluation of indoor air quality was relatively higher (3.97) than that of outdoor air quality (3.25). However, on a scale of 1–9, average rating by residents of home air quality was 6.08—indicating a Healthy perception—while average neighbourhood and city air quality ratings were 4.33 and 3.83 respectively. However, inspectors were found to overestimate air quality across the various domains as well. But relative to inspectors, residents' ratings leaned more positively for home air quality ($p=0.001$). Therefore, Residents also tended to evaluate indoor air quality as healthier than what the data showed, and more positive relative to inspectors. A 'Home Halo Effect' is measured by quantifying the difference between overestimation of indoor air quality in polluted households between residents and inspectors. Residents overestimate air quality by 80% while inspectors only by 70%, meaning a 10% halo effect according to the authors' definition, whereby the percentages represent the percentage difference between average subjective evaluation and the objective measurement.

Building on the questionnaire data, Hofflinger (2019) sought to further establish the mechanisms that determine public perceptions of air quality as well as prove the prevalence of the halo effect over other related elements. Several independent factors—household income, education level, age, presence of respiratory ailments, number of wood-burning stoves, and whether they live in areas where wood-burning stove usage is restricted—were weighed against data collected on air quality evaluation in a mathematical model to investigate four dependent variables: perception of air quality in the home, neighbourhood, and the city, and the home halo effect. Results displayed a clear pattern of better subjective evaluation of air quality at home, as compared to the city or the neighbourhood—61% of respondents rated home air quality better than city, while only 7% thought it was worse. The mean value of the halo effect variable was positive, confirming a general trend of perceiving home air as cleaner than it is. These results are consistent with the similar spatial cognitive bias, the “neighbourhood halo effect” (Bickerstaff and Walker, 2001), termed in accordance with evidence found that individuals generally perceived communities in close proximity to theirs to be less polluted than others. (Rankin, 1969). Both studies support the finding that people rate environmental conditions in their immediate surroundings as significantly better than broader or more distant areas.

Since the significance of the halo effect in public perception of air quality has been steadily supported well established through quantitative studies, applying these insights to urban policy can potentially lead about to improvements. After all, the public's understanding of and response to air pollution can have significant implications for the

Formatted: Default Paragraph Font, Font: (Default)
Lora, 11 pt, Font color: Custom Color(RGB(34,34,34)),
Highlight

effectiveness of government interventions aimed at fostering the health and wellbeing of urban residents (Hofflinger et al., 2019). We ~~could even go as far to say that~~ might even argue that the success of mitigation policies hinge is largely dependent on a thorough understanding of people's attitudes towards indoor and outdoor air quality.

~~I suggest~~ A proposed application is that governments can provide subsidies or tax rebates to real estate projects ~~in fashionable neighbourhoods~~ that meet a high threshold of integrated air-purifying green ~~spaces into their buildings~~ infrastructure, targeting this support in neighbourhoods that already carry a strong positive reputation. By doing that, the government can incentivise a rise in green infrastructure in trendy urban areas. This would link pollution-reducing infrastructure with places that are seen as attractive; modern, or affluent, leveraging the halo-effect to make sustainable features desirable status symbols. Residents would then begin to associate environmental improvement with aesthetic value, influencing them to incorporate it into their own home designs ~~As residents and outside observers already hold a favourable global impression of such neighbourhoods, this positive impression should colour their evaluation of the air-purifying infrastructure introduced. This means the green design features are more likely to be judged favourably, as more effective, more desirable, or more worth emulating.~~ However, it is important to note potential equity concerns with such a scheme: fashionable neighbourhoods are typically more affluent, so subsidy provision may neglect less well-off neighbourhoods that lack the additional income to integrate sustainable design.

At the community level, local environmental agencies could partner with platforms to offer verified "Green-Status" badges on their user profiles, earned by documenting air-friendly actions (e.g. using public transport, installing an air-quality monitor). This allows environmental sustainability to be associated with social credibility ~~(Casillia and Hickey, 2016) and be cast under its~~ in its positive halo. Since ~~people individuals are inclined towards accumulating love to gain~~ social capital ~~(Weiler et al., 2022)~~ and follow social trends, being eco-conscious will come to be seen as ~~the "in" thing~~ fashionable, encouraging widespread adoption of green behaviours through social proof. ~~That said, we should acknowledge the issues of performative behaviour and unequal platform access, which could compromise the initiative's effectiveness in achieving real and widespread behavioural change.~~

V. Confirmation Bias

While ~~most of us many~~ may consider ~~ourselves themselves~~ open to learning and accepting ~~new concepts and points of view~~ new ideas, ~~what we may not realise is our~~ individuals tend to tendency to cherry-pick and interpret information in a manner that adheres to ~~our their~~ existing beliefs, ~~and rejecting or discounting~~ contradictory information ~~that contradicts~~

our beliefs. This phenomenon is attributed to “confirmation bias”, which, Nickerson (1998, p. 1) defines as “the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand”.

Following in the footsteps of philosopher Karl Popper (1959), who believed hypotheses should be falsified instead of confirmed, Wason (1960) sought to experimentally demonstrate the existence of a confirmation bias. Participants were given three numbers “2-4-6” and told that the triple satisfies a predetermined rule. Each round, they were then asked to construct a triple they believed fulfilled the rule as well, and subsequently feedback to whether the rule was fulfilled was given. When ready, participants were to state their assumed rule, then asked to infer the rule by proposing their own triples and receiving feedback after each attempt. Participants generally behaved in the following way: typically they formed a hypothesis on the rule (e.g. “consecutive even numbers”) then tested their hypothesis by only proposing triples that satisfied the rule (e.g. “4-6-8”, “6-8-10”, “20-22-24”). As the feedback was always positive, their mistaken assumptions were reinforced till they ultimately presented their incorrect hypotheses after several rounds of testing, they were told they were wrong. In actuality, the rule was simply “ascending numbers”. The methods participants employed to test their hypothesis bore affirmative evidence despite their assumptions being wrong. Their tendency to seek supporting rather than disconfirming evidence, exhibiting their use of a confirmation bias to arrive at their eventual conclusions (Oswald and Grosjean, 2004).

Expanding on this, Lord, Lee-Ross, and Lepper (1979) examine how confirmation biases go further to increase belief polarisation—rather than decrease or leave it unaffected—when mixed or inconclusive findings are assimilated by proponents of opposite viewpoints.²⁵ Their study explores views on capital punishment between two groups of subjects recruited such that they were evenly split between being pro-deterrence and anti-deterrence. The study consists of two stages. In the first stage, subjects were given a brief overview of a study’s results, then asked to rate their change in attitudes toward capital punishment. In the second stage, subjects were provided with a detailed presentation of the procedure, results, and critiques, and again asked to evaluate its impact on their attitudes. This was then repeated with the second study. The following results are presented as percentage point changes. The first stage’s results reported proponents shifting their views 3.75% towards becoming more in favour of capital punishment, whereas opponents shifted 3.13% towards being more opposed. Notably, the divergence widened in the second stage, with the beliefs of proponents and opponents strengthening in opposite directions by 9.38% and 10.6% respectively, despite substantial evidence being provided for both supporting and opposing views. Such data strongly supports that the

view that mixed data will can lead to increased polarization polarisation rather than to uncertainty and moderation.

The confirmation bias is highly relevant in the context of air pollution, lying in the fact that airAir pollution has been a recognised problem for millennia (Fowler et al., 2020). Entrenched beliefs about atmospheric pollution, often solidified as social stigmas, can persist even when their original scientific basis becomes outdated (King, 2014; Schwartz, 2004). This persistence helps explain the limited impact of modern information campaigns, which seems to prompt limited environmental action (Ramirez et al., 2019).

Previous research, albeit sparse, substantiates is consistent in supporting the correlation between the confirmation bias and air pollution. Schwartz (2004) took an observational approach by comparing objective air pollution rates in the United States against subjective evaluation by citizens.²⁷ He cited public polls which found that the majority of Americans believed air pollution has gotten worse in the past 10 years (66%) and would continue to get worse (67%). On a state level, 78% of Californians were reported to believe the state has hardly made progress in solving environmental problems and around 80% of New Yorkers rate air pollution as a “very serious” or “somewhat serious” problem, as do 77% of Texans. Ironically enoughIn fact, actual ozone level data reflect that the US has made dramatic progress in reducing air pollution over the last few decades, and most American cities now enjoy relatively good better air quality than in the past. The nation's monitoring stations registered ozone in excess of federal standards more than 12 times per year in the early 1980s, while the end of 2002 saw a fall to 4 exceedances per year. The considerable improvement in air quality should be perceptible, yet there remains a large disconnect between perception and reality. Based on the evidence presented, it is inferentially likely that the presence of confirmation bias prompts led individuals to cling to historically negative viewpoints of American air quality and overlook objective improvements, impeding their ability to perceive environmental impacts generated by measures put in place to mitigate air pollution. However, the possibility of alternative explanations such as media exposure, uneven local improvement, and distrust of official data is to be acknowledged. Although the above data is historical, the long period over which the study was conducted offers insight into the enduring nature of confirmation bias, which is relevant to current policy design.

An observational study by King (2014) probes into cases where confirmation bias contributes to false perceptions of air pollution due to influence from preconceived notions tied to socioeconomic levels.⁴ The study is set in Chicago, where—in contrast to the national trend—non-Hispanic Blacks African Americans are slightly less likely than other groups to live near toxic-releasing facilities and Hispanics are more likely to be

exposed—consistent with a national tendency for higher environmental risk near concentrations of immigrants. Key results are as follows:

Air quality rating	Poverty rate of raters	Proportion of African-American raters	Proportion of Hispanic raters	Additional notes
Poor	21.5%	37.9%	33.6%	46.9% of residents without a high school diploma rated air poorly; 28.4% of college graduates did so.
Excellent	10.9%	22.5%	20.9%	=

Table 1: Summary of statistics for socioeconomic groups by proportion by air quality rating (adapted from King, 2014, p. 21)

residents who rated their air quality as “Poor” lived in neighbourhoods where, on average, 21.5% of families were in poverty; conversely, those who rated their air as “Excellent” lived in area with only 10.9% poverty; “Poor” raters also lived in areas that were 37.9% Black and 33.6% Hispanic on average, whereas “Excellent” raters lived in areas that were only 22.5% Black and 20.9% Hispanic; 46.9% of residents with less than a high school education rated their air poorly, while only 28.4% of those with a college degree did the same. Despite the stark perceptual difference, the general difference in objective pollution levels between people who reported “Excellent” air and “Poor” air is actually small. Thus, social cues (in this case, racial demographics) emerge as the a likely key predictors of air quality, a finding consistent with a neighborhood stigma interpretation. However, given these are correlations from observational data, there can be several potential drivers of these associations including news coverage patterns or literacy levels.

King suggests a “neighbourhood stigma” theory wherein “once a place acquires a negative reputation, respondents will tend to rely heavily on that evaluation when inferring neighborhood characteristics, particularly in the absence of domain-specific information or experience”. While place reputation appears to be affected by individual bias dependent on of racial/ethnic phenotypes racial and ethnic composition in Chicago, other considerations may be more salient in other areas. This relates back to confirmation bias by demonstrating how people are inclined to make judgements about air quality that align with preconceived social beliefs.

Fully grasping the implications and applications of confirmation bias can enable policy makers to not only alleviate its negative effects but harness it to positively influence public

Formatted Table

Formatted: Centered

Formatted: Font: (Default) Lora, 11 pt

behaviour. ~~Here I propose~~ Two communication strategies are suggested for regulators ~~to~~ can adopt to aid in leverage mitigating on confirmation bias barriers in the battle against air pollution.

The first strategy uses confirmation bias to link environmental messaging to identity and values. drawing on research into political moral framing in environmental action by Hurst et al. (2020) and Boso et al. (2025). To do this, policymakers can frame pollution issues in a manner that aligns with different groups' values. This plays into on their confirmation biases, the presence of which is consistent with evidence from Ditto et al.'s (2017) meta-analysis showing that political partisans reject information that counters their pre-existing beliefs, and could potentially make them more receptive to proposed environmental actions. This can be particularly useful for garnering conservative support for environmental policy, given their relative distrust of climate science (Washburn et al., 2017) and general refutation of anthropogenic climate change (Ballew et al., 2019). For instance, instance, conservatism as an ideology favours a free market economy, a family health and Thus, an economic economy efficiency-oriented narrative that emphasises economic benefits from efficiency gains or economic costs of environmental disasters and resource depletion can be can be channelled towards conservatives while climate justice and intergenerational responsibility can be emphasized to progressives. For a conservative audience, messaging could highlight how reducing air pollution protects children's health and lowers long-term healthcare costs. For a progressive audience, there could be a greater focus on the disproportionate impact poor air quality has on marginalised communities. more persuasive than traditional information sharing that describes the effects of climate change.

The second one involves curating source networks and information channels that hold sway over different groups. This approach leverages confirmation bias by strategically delivering the value-aligned messages from the first strategy through sources and channels the target audience already trusts and seeks out, drawing on Shen et al.'s (2023) evidence for the influence of trusted community-based organisations in public health messaging. For a conservative audience, this means partnering with and amplifying messages through faith leaders, military veterans' groups, and centre-right media personalities. For a progressive audience, effective channels include social justice organisations, educational institutions, and progressive media outlets. The above suggestions rely on a simplification of political groups and acknowledge variation within groups.

However, the ethics of using confirmation bias should be considered before any real-world integration, for there is a thin line between value-aligned communication and manipulative targeting. To ensure ethical integrity, all information conveyed should remain factually correct and scientifically consistent, with only the emphasis in messaging varying. This is

to further prevent unintentionally deepening divisions by allowing different groups to receive substantially different narratives (content-wise) about the same issue.

VI. Loss Aversion

A coin is tossed and the following question is posed to you: "If it's tails, you lose \$10. How much would you have to gain on winning in order for this gamble to be acceptable to you?" Was your answer around \$20? Kahneman illustrates a standard coin-toss gamble experiment in *Thinking Fast and Slow* (2011). He asks participants how much they would need to gain if they stood to lose \$10 on tails. The typical response was approximately \$20. That was the answer most of Kahneman's students gave him. From this demonstration, we can plainly see suggesting a tendency that people to place twice as much more value on losing money as opposed to gaining money. This is —a textbook example of the loss aversion hypothesis. (Kahneman, 2011)

Formatted: Font: Italic

The loss aversion hypothesis was first introduced by Kahneman and Tversky (1979) as an essential feature of prospect theory, which suggests people make decisions based on perceived value relative to a reference point instead of a static utility. The hypothesis is that "the response to losses is more extreme than the response to gains. The displeasure associated with losing a sum of money is generally greater than the pleasure associated with winning the same amount" (Kahneman and Tversky, 1981, Pg. 454). A consequence of loss aversion is that it can lead to inconsistencies in risk preferences, i.e. the degree to which an individual prefers a sure amount to a gamble for any arbitrary risk. It is important to note the distinction between loss and risk aversion, where the latter refers specifically to a preference for the certain outcome over a gamble with the same expected value (Menezes and Hanson, 1970). Essentially, Loss aversion can therefore elicit individuals' risk-averse behaviour, and such behaviour may be used to infer loss aversion. An individual is risk-averse if they prefer the sure amount equal to the expected value of the gamble to the gamble, and risk-seeking if they prefer the latter. (Menezes and Hanson, 1970)

Karle et al. (2013)⁵ demonstrated the mechanisms of the heuristic in an experimental setting using choices among different lotteries.³² They gave subjects two series of choices, with 6 choices each, where they had to choose between lotteries and sure payments. For Series A, subjects had to make 6 choices between a lottery with a 50% chance of winning 1 Euro and a 50% chance of winning nothing, and a sure payment of S. S was either 10, 20, 30, 40, 50, or 60 Eurocents. This choice purely elicits their risk preferences without involving loss aversion as it only comprises gains. A risk-neutral individual would be indifferent between the lottery (which has an expected value of 50 Eurocents) and a sure payment of 50 Eurocents. Subjects who choose the lottery over sure payments below of 50

Eurocents ~~and above over the lottery~~ are risk-seekers~~averters~~, while those who choose ~~the lottery over~~ sure payments ~~of above~~ 50 Eurocents ~~and below over the lottery~~ are risk-~~averters~~~~seekers~~. In the experiment, ~~only 76.7%~~23.3% of participants picked ~~the lottery over~~ sure payments ~~above the lottery's expected value a sure payment that is below or equal to~~ 50 cents (risk-~~averters~~~~risk-seekers~~), ~~showing that most participants were either risk-~~neutral or risk-averse. -

In series B, subjects had to make ~~six seven~~ choices between a lottery that gave a $\frac{1}{3}$ chance of winning €1 and a $\frac{2}{3}$ chance of losing R Eurocents, and a sure payment of zero. R was either 0, 10, 20, 30, 50, 70, or 100 Eurocents. This choice is where loss aversion comes into play. Going by the same logic in Series A, a risk-neutral individual would be indifferent between a lottery of expected value zero (when R=33) and the sure payment. ~~Meanwhile,~~ ~~a~~An individual that ~~additionally, displays exhibits~~ loss aversion, ~~and thus risk-aversion,~~ would be willing to accept a payment of zero more often for any given R. In the results of the experiment, 16.8% of participants make a choice that is more risk averse compared to what they did in Series A and therefore display a change in behaviour due to loss aversion.

Pollution links closely to loss aversion in terms of environmental action by individuals or groups. Whether or not to undertake a pro-environment action is a choice individuals have to make. That decision making process likely involves weighing private benefit derived against private cost personally incurred. ~~Previous research done on the topic has shown that when faced with two possible choices, one energy-efficient and the other not, people will only choose the energy-efficient item if the payback time for the energy saved is very short~~ (Kollmuss and Agyeman ~~write that "if the person decides between two possible items, one energy-efficient and the other not, he or she will only choose the energy-efficient item if the payback time for the energy saved is very short"~~ - (2002, p. 249). This is consistent with loss-averse behaviour as ~~it suggests a greater valuation of they value the loss of their time more than possible the environmental benefit~~. Moreover, since partaking in pro-environmental behaviour often requires a lifestyle change and a loss of the familiarity and safety in one's usual routine or incurs implementation costs, loss-averse individuals are likely to opt for the non-environmental option. ~~A limitation of this analysis is that it does not disentangle loss aversion from other alternative explanations such as heterogeneity in time preferences and liquidity constraints.~~

Bartczak et al. (2015) explores loss aversion in the environmental domain via a within-subject field study examining choices over a risky environmental good, wildfires in Poland.⁵ The study adapts Tanaka et al.'s (2010) experimental lottery design where ~~Polish individuals~~participants chose among lottery pairs framed in terms of environmental outcomes of programs to reduce the risk of forest wildfires rather than monetary payoffs. ~~are presented with three series of lottery pairs and asked to choose one lottery for each~~

Formatted: Font: Lora, 11 pt

Formatted: Font: Lora, 11 pt

Formatted: Font: Lora, 11 pt

Formatted: Font: Lora, 11 pt

Formatted: Font: Lora, 11 pt

Formatted: Font: Lora, 11 pt

Formatted: Font: Lora, 11 pt

pair.³⁴ Respondents are asked to make a choice between programs that, instead of financial rewards, presented environmental effects of programs to reduce the risk of forest wildfires. In Treatment 1 and 2, the first two treatments, they which involved only gains (choice between programs with varying success rates) are told that there are two potential programs to further reduce the risk of forest fires which carry different success rates, and results are described as the annual average reduction of burned forest area in Poland and presented as uncertain. $\approx 75\%$ of respondents were risk-averse. Simply put, the first two treatments involve choosing between only gains. As detailed above in Karle et al.'s study, this is used to determine risk preference independent of loss aversion. In Treatment 3 (the core treatment) the third treatment, which introduced the possibility of losses through greater forest destruction (choice between programs with varying success and failure rates), respondents are asked to choose between programmes which would end in either success or failure, where more forests would be destroyed than currently. As there is now an element of loss, subjects' loss aversion can be elicited. Results from Treatment 1 and 2 indicated 75% of respondents were risk-averse. However, in Treatment 3 which introduces losses, the majority ($\approx 85\%$) of respondents exhibited risk-seeking behaviour instead. $\approx 85\%$ of respondents were risk-seeking. This suggests that the introduction of losses substantially altered choice behaviour. The paper thus provides conclusive evidence for loss aversion in the environmental context.

More recently, Knobloch et al.'s (2019) paper displays loss aversion's widespread applicability to the global energy model, using residential heating technology uptake as a case study.³⁵ The sector was chosen as decisions are overwhelmingly made by consumers, but not predominantly influenced by other interfering cultural factors (Michelsen et al., 2013). Additionally, the sector accounts for the largest share of direct greenhouse gas emissions by households. It is a simulation ~~n-observational~~ study that performs analyses at both a global level and the level of individual countries, simulating future heating technology diffusion across 59 world regions.

The study employs a model that simulates consumer decision-making and weighs benefits (convenience, environmental benefits, etc.) and costs (upfront costs, operating costs, etc.). The model is run twice, with default loss aversion parameters and without. The model is used to compare the efficacy of four decarbonization policy scenarios with and without the effect of consumer loss aversion. These policy scenarios involve different levels of residential carbon taxes and subsidy payments. In this case, a carbon tax acts as the "gain" element as it increases the size of relative gains from switching to renewables by reducing energy costs, while subsidies act as the "loss" element as they reduce the size of relative losses. The aim is to assess how loss aversion influences model projections of technological change and policy effectiveness, especially as current energy models often assume optimal consumer decisions, which might lead to overly optimistic expectations of climate policy

efficacy. Results indicate that loss-averse decision-makers are less likely to adopt energy-efficient technologies, as higher upfront costs are often perceived as losses with a stronger impact than future energy savings (evaluated as gains). This asymmetric valuation results in a "conservative shift" where preferences for current technologies are strengthened, reducing uptake of newer technologies.

The model further shows that subsidies were relatively more effective in reducing emission levels than carbon taxes as they directly address the pain point for loss-averse consumers—high investment costs for renewables, while taxes do not. On the other hand, projections show that a carbon tax of 200 €/tCO₂ while considering loss aversion and a carbon tax of 100 €/tCO₂ without considering loss aversion achieve a similar reduction in overall emission levels. Thus, taxes require more drastic implementation, which is likely to be received poorly by the public. In short, the model tells us indicates that a tax rate about we need double the tax that of ideal projections (without loss aversion) is required to attain similar outcomes to see plausible results comparable to idealistic projections (without loss aversion).

These differences in outcomes suggest that consumer behaviour, specifically loss aversion, has a profound quantitative impact on the feasibility and required intensity of climate policies.

Having proven how loss aversion pertains to environmental pollution, policymakers can tap on this heuristic to maximise the effectiveness of policies aimed at reducing pollution. Taken together, these findings suggest that loss aversion may play an important role in environmental pollution and inaction. Thus, policymakers may be able to use this heuristic to increase the effectiveness of pro-environmental policy. While loss aversion is often discussed in the context of how it can hinder the effectiveness of policies, there are now several policies and interventions that have sought to actively "make use of" this behavioral trait to achieve desired outcomes. Current and previous approaches often involve framing choices to emphasize potential losses, thereby leveraging individuals' stronger aversion to losing something compared to gaining an equivalent amount.

A notable application is in the design of incentives for healthy behaviours, where individuals are offered "insurance" (a sum of money) but told they will lose it if they do not achieve a certain healthy behaviour (e.g. attending an appointment or making a healthy choice). The prospect of losing this already "allocated" amount can be motivating and has been shown to encourage high rates of health behaviour uptake (Meeker et al., 2015).

This strategy can be adapted for pollution reduction in various ways, particularly because the considerations underpinning pro-health and pro-environmental behaviour are similar.

Both rely on long-term commitment for delayed gains, and overlap in terms of sustainable lifestyle habits, such as setting household carbon emission thresholds. A “green bonus” can be granted to households at the start of each month (e.g. credit on their utility bill) but lose part or all of it at the end of the month if they exceed the given emission threshold. This creates a stronger motivation to adopt greener appliances and habits than a standard rebate system: by exploiting loss aversion, as households are more sensitive to losing a granted credit than to forgoing an equivalent unheld gain. However, this design raises concerns about one uniform threshold. Large households may be disadvantaged as their consumption is structurally higher; lower-income households often have little room to cut, while higher-income households have more slack and thus an easier path to the bonus, risking a regressive outcome; and renters have limited control over efficiency, insulation, and heating decisions. Thus, adjusted thresholds to accommodate varying baselines are necessary to reward improvement rather than penalise pre-existing consumption needs.

Alternatively, ~~c~~Commuters could begin each week or month with a “streak bonus” (e.g. 10% discount on fares) that diminishes if they fall short of their weekly quota. In other words, missing a day of commuting via public transport would constitute a penalty (e.g. - 0.5% off the discount). However, a fixed quota assumes consistent transit access, which does not hold for commuters with disabilities limiting service use, or those in rural or low-density areas where frequency and coverage are constrained by infrastructure, punishing accessibility limits. The scheme also risks disproportionately affecting low-income commuters who lose their discount due to illness, caregiving, irregular work hours, penalising them more harshly (since discount makes up a larger proportion of their transport budget) instead of rightfully supporting them.

These examples ~~serve may help into~~ nudging individuals to ~~more willingly~~ adopt greener lifestyle habits ~~more readily~~.

~~V~~**VII. Discussion**

~~Up tntill~~ Until this point, each heuristic has been examined independently. However, in practice they do not operate in isolation. Although not limited to these three biases, we see evidence of a ~~possible perpetual~~ compounding effect ~~even just~~ within the three heuristics examined in this paper, demonstrating a mutually reinforcing relationship not yet explored in previously cited studies. This section proposes a ~~possible-potential~~ account of ~~how~~ such an interaction might unfold.

One way of conceptualising this interaction is as a sequential cycle, in which each heuristic builds on the perceptual conditions created by the preceding one. This is not offered as a

Formatted: Font: (Default) Lora, 14 pt, Bold, Font color: Text 1

definitive model but as a plausible account that broadly draws on is broadly consistent with the evidence reviewed in earlier sections.

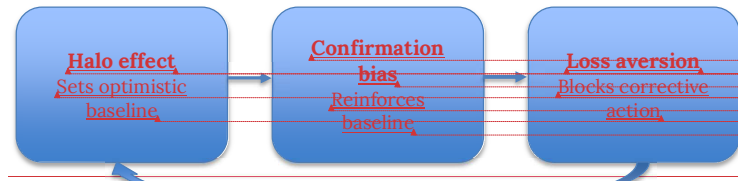


Figure 1: Proposed framework

The halo effect may be understood as the stimulus that establishes the individual's baseline perception of local air quality. The theoretical basis for placing it first in the sequence is that, by definition, the halo effect is a phenomenon of global evaluation—it operates at the level of overall impression, shaping how an individual feels about a place or environment before any deliberate assessment of its specific attributes has taken place (Nisbett and Wilson, 1977). However, alternative sequencing remains possible as prior beliefs or loss concerns can precede place attachment. As demonstrated by Hofflinger et al. (2019) and Boso et al. (2020), (positive) emotional attachment to one's home and neighbourhood generates a global positive evaluation that colours assessment of individual attributes, including air quality, in a more favourable direction than objective data warrants. This initial perceptual anchor is not consciously constructed; it emerges automatically from the same System 1 processing that underlies heuristic reasoning more broadly. The critical point is that once established, this baseline does not simply sit passively in the individual's mind; it actively colours how subsequent information is processed.

This is where the confirmation bias plausibly enters, operating as a filter through which subsequent information is screened rather than processed neutrally. Having formed an overly optimistic baseline perception, the individual is now predisposed to filter attend selectively to incoming information in a manner that preserve rather than test it that confirms it. A hypothetical example is that On a clear, sunny day, the absence of visible haze may be registered as evidence of clean air; a health advisory may be discounted as overstated; a neighbour's complaint about air quality may be attributed to hypersensitivity. As Schwartz (2004) and King (2014) demonstrate, such selective interpretation this filtering process is capable of sustaining misperceptions across years and even decades, insulating individuals from the corrective influence of objective environmental measures. In this case, the halo-derived baseline provides the very belief that the confirmation bias then works to protect and entrench.

- Formatted: Centered
- Formatted: Font: Lora, 11 pt, Bold
- Formatted: Font: Lora, 11 pt, Bold
- Formatted: Font: Lora, 11 pt
- Formatted: Font: Lora, 11 pt, Bold
- Formatted: Font: Lora, 11 pt
- Formatted: Font: Lora, 11 pt
- Formatted: Font: Lora, 10 pt
- Formatted: Font: Lora, 10 pt
- Formatted: Font: Lora, 10 pt
- Formatted: Font: Lora, 11 pt, English (United States)
- Formatted: Font: Lora, 10 pt, English (United States)
- Formatted: Font: Lora, 11 pt, English (United States)
- Formatted: Centered

The third stage introduces loss aversion as a barrier to behavioural correction cases where the first two stages are partially overcome—where an individual does come to accept that local air quality is worse than they had assumed—the translation of updated perception into action faces a further obstacle. As Knobloch et al. (2019) and Bartczak et al., (2015) elucidate, the upfront costs, lifestyle disruptions, and sacrifices of familiarity associated with pro-environmental behaviour are experienced as losses, weighted disproportionately against the future gains of cleaner air. Loss aversion, understood as such, functions are as the final barrier in the cycle. It is the mechanism by which even a corrected-changed perception fails to produce a corrected-changed behaviour. The individual may acknowledge the problem and yet remain inert.

Taken together, this proposed sequence suggests a possible dynamic in which the halo effect generates the misconception, confirmation bias reinforces it, and loss aversion prevents actionable progress even when negative perceptions are dislodged. Future work could test this hypothesised causal chain using a full factorial 3x3 experimental design where each treatment dimension is one of the heuristics in the causal sequence. While the precise causal ordering directionality (may be multidirectional rather than strictly unidirectional) proposed above remains an open empirical question, the critical takeaway is that when multiple heuristics operate simultaneously in the same cognitive environment, the result may be qualitatively different from, and more intractable than the basic sum of their individual effects.

Two of the cases discussed in earlier sections are worth revisiting through this integrative lens, as they are suggestive —if not conclusive—of the kind of compounding dynamic described above.

The first is the wood-burning stove context in Southern Chile examined by Boso et al. (2020). Residents in Temuco and Padre Las Casas systematically overestimated indoor air quality relative to both sensor measurements and inspector evaluations—a pattern consistent with the home halo effect. Yet this misperception persisted despite aggressive government intervention, including refurbishment subsidies and prohibitions on firewood use, suggesting that objective information and material incentives that should rationally tip the scale in favour of actionable change were insufficient to shift behaviour. This resistance is arguably consistent with confirmation bias entrenching the favourable baseline perception against contradictory policy messaging, while loss aversion simultaneously raised the psychological cost of abandoning familiar heating practices. An important caveat is that their data is purely correlational so alternative structural explanations are plausible.

Formatted: Not Highlight

The second case is King's (2014) study of air quality perception in Chicago. This illustrates that the compounding dynamic need not operate exclusively in the direction of over-optimism. Here, neighbourhood stigma (negative halo effect), generated a pessimistic baseline perception of ~~air~~ quality in predominantly Black and Hispanic neighbourhoods, despite the objective difference in pollution levels between high- and low-rated areas being small. Confirmation bias then plausibly entrenched this stigma, causing residents to interpret ambiguous environmental cues through a lens of pre-existing negative expectations. An open research question is that if loss aversion may further have discouraged investment in local environmental improvement, with residents perceiving the costs of action as disproportionate given the low perceived likelihood of meaningful change. The cycle, in this case, runs in the opposite direction—producing systematic underestimation of air quality rather than overestimation—but the underlying compounding logic is arguably the same.

These two cases do not constitute proof of the proposed interaction. They are better understood as illustrations of contexts in which such an interaction is a plausible contributing explanation for outcomes that individual heuristics alone do not fully account for.

If the compounding dynamic proposed here has merit, it carries meaningful implications for the design of pollution-related policy. Interventions targeting only one heuristic at a time may find their effectiveness partially undermined by the others operating in the background. For example, a communication campaign that successfully challenges a community's confirmation bias, for instance, may still fail to produce behavioural change if loss aversion remains unaddressed—the individual may update their beliefs without updating their behaviour, consistent with the intention-behaviour gap in medical literature (Faries, 2016).

This suggests that layered policy approaches—ones that address the emotional baseline, the information environment, and the incentives structure in coordinated fashion—may be more effective than single-lever interventions, even if they are considerably more complex to design and evaluate. The individual policy proposals outlined in Sections II, III, and IV might therefore be understood not merely as standalone recommendations but as potentially complementary components of a broader behavioural strategy ~~but as a potentially complementary component of a broader behavioural strategy~~: reshaping the halo first to establish a more accurate perceptual baseline, ~~then~~ delivering value-aligned messaging to work with rather than against confirmation bias, then structuring policy design to align with loss incentives.

This document was
truncated.

To convert this file
completely, without the
online limits, use the desktop
version of Total PDF

Converter:

<https://www.coolutils.com/TotalPDFConverter>



COOLUtils.com
file converters



Response to Reviewers of The Psychology of Pollution: Analysing How Heuristics Affect Perception of Air Quality and Policy Design

I sincerely thank both reviewers for their thorough and constructive comments. In the following, I address each comment point-by-point and indicate where the corresponding change was made in the revised manuscript. I have separated responses to each reviewer for ease of re-review. Reviewers' comments are copy-and-pasted and listed as bullet points, with responses indented beneath each one.

Reviewer 1's comments

Major Comments:

- In the section on the halo event, the proposed policy implications are creative but the mechanism requires clarification. The intervention relies less on the traditional halo effect (i.e., existing positive impressions influence judgments of unrelated attributes) and more on social signaling/identity-based motivation, where sustainability becomes associated with status and desirability. A more appropriate example of the halo effect would be for example that affluent neighborhoods are assumed to have cleaner air/better environmental quality even when it is not the case. I suggest that the author considers this distinction and alters the proposal to be more in line with the halo effect or provide evidence supporting the assumption that exposure to green infrastructure in prestigious places leads to broader behavioral adoption. Additionally, there are ethical/equity concerns to consider here. Investments in green infrastructure in affluent neighborhoods would disproportionately benefit already advantaged communities – and in fact this is already happening today. Consider reframing the intervention to ensure these behavioral insights are used to increase adoption of sustainable infrastructure across different communities, especially those experiencing the greatest pollution burdens. At minimum, consider mentioning this nuance.
 - The proposal was rewritten so it operates more directly through the traditional halo effect: subsidies are targeted at neighbourhoods that already carry a strong positive reputation, so the pre-existing favourable impression of the area colours residents' and outsiders' evaluation of the new green infrastructure itself. An explicit equity caveat was added noting that fashionable neighbourhoods are typically more affluent, so additional subsidy provision risks neglecting less well-off neighbourhoods. A parallel caveat on performative behaviour and unequal platform access was added to the Green-Status badge proposal. [Section IV (Halo Effect), real-estate subsidy and Green-Status badge paragraphs]
- For the section on confirmation bias, the author gives a strong overview and does a good job connecting foundational psych work to perceptions of the environment. However, be careful in distinguishing between evidence of misperception vs confirmation bias. E.g., a discrepancy between objective measures and subjective perceptions of air quality doesn't necessarily show confirmation bias, as there are many alternative explanations such as differences in personal experience, distrust in monitoring systems/science, media exposure, changing expectations/standards, or attention. To more directly support the

confirmation bias argument, consider focusing more on evidence that shows people selectively seek out, interpret, or remember information in ways that reinforce pre-existing beliefs. Additionally, I'd stay away from using phrases like "leveraging confirmation bias" as it is typically thought to be a negative barrier that stops people from processing info accurately. You don't want to seem as if you suggest exploiting confirmation bias but rather that policymakers can work with existing values/identities to improve receptiveness to information. Relatedly, the author may want to acknowledge possible risks of targeting messaging based on community as it may unintentionally deepen divisions by allowing different groups to receive substantially different narratives about the same issue. Consider discussing how identity-based communication can be used ethically while still maintaining scientific consistency.

- Added explicit acknowledgement that alternative explanations (media exposure, uneven local improvement, distrust of official data) could account for the perception-reality gap, and reframed the confirmation-bias interpretation as "inferentially likely" rather than proven. Replaced "leveraging"-style phrasing with more neutral language. Added a new closing paragraph distinguishing ethical value-aligned communication from manipulative targeting, requiring all messaging to remain factually correct and scientifically consistent, and noting the risk of deepening division through divergent narratives. [Section V (Confirmation Bias), paragraphs following the Schwartz/King discussion and the final paragraph of the section]
- The loss aversion section makes an interesting argument about how loss aversion affects environmental behavior, but its connection to the paper's broader focus of air pollution perception is less clear. Unlike halo effect and confirmation bias sections, this section primarily addresses behavioral responses rather than how people perceive pollution risks. Consider explicitly explaining whether loss aversion acts as a barrier after perception has already changed (i.e., people recognize pollution as a problem but avoid costly solutions), rather than presenting it as another mechanism shaping perception itself. Also, the proposed applications to environmental policy are interesting, but should discuss possible limitations of these approaches. For example, penalties or forfeitable incentives may disproportionately affect households with fewer resources or limited ability to adopt greener technologies
 - The discussion now explicitly positions loss aversion as the third, behavioural stage of the model (the mechanism by which a "changed perception" fails to produce a "changed behaviour") rather than a perception-shaping mechanism in its own right. Both the green-bonus and commuter streak-bonus proposals were expanded to flag distributional risks (household size/income effects, renter control over efficiency measures, disability and rural transit access, and disproportionate impact on low-income commuters), with adjusted thresholds to mitigate regressive outcomes. [Section VII (Discussion), third-stage paragraph; Section VI (Loss Aversion), green-bonus and streak-bonus paragraphs]
- I think this proposed interaction model is the most original aspect of the manuscript, but requires more careful framing. The studies reviewed earlier provide evidence that each heuristic may influence perceptions/behavior independently, but don't show these biases operate sequentially or causally in the order proposed by the author. For example, the halo

effect does not necessarily need to precede confirmation bias, and loss aversion may influence whether individuals seek or interpret information in the first place. I suggest softening the framing of this section to frame it as a conceptual framework or hypothesis rather than a proposed sequence supported by existing evidence. Also, you might consider strengthening this section by outlining specific predictions that future empirical work could test. Alternatively, consider if these are better conceptualized as interacting factors rather than fixed sequence. A cyclical model may better capture that they may reinforce each other in multiple directions instead of occurring in order.

- The discussion now explicitly frames the sequence as "not...a definitive model but a plausible account," states that "alternative sequencing remains possible," and notes that directionality "may be multidirectional rather than strictly unidirectional." A new closing paragraph proposes a factorial experimental design manipulating the heuristics together and states specific, testable predictions for future empirical work. A schematic (Figure 1, "Proposed framework") was also added to visually mark the model as a proposal. [Section VII (Discussion), opening and final two paragraphs]
- The conclusion is generally strong and appropriately synthesizes the paper, but overreaches on some claims (similarly to earlier sections). I suggest softening some of the claims to better reflect the strength of the evidence presented by using language such as "may substantially influence" or "can contribute to" rather than definitive, established mechanisms.
 - Causal language was softened throughout the Conclusion ("has provided compelling evidence" changed to "has suggested meaningful influence"; "explains" changed to "may help to explain"; "leads" change to "offers a plausible mechanism"), and an explicit caveat was added noting that because the underlying studies are observational, confirmation bias cannot be isolated as the cause of persistent perceptions and the association should be read as suggestive rather than demonstrated. [Section VIII (Conclusion), paragraphs 1–3]

Minor Comments:

- "Economic costs amounting to trillions of dollars—eroding national GDPs and straining healthcare systems—are starkly documented." □ no citation
 - Rewritten as a complete sentence ("...are stark documentations of pollution-induced economic instability."), sitting within the paragraph anchored by the surrounding WHO (2016) citation. [Section I (Introduction), paragraph 1]
- "These deviations from neoclassical economics are instances of what is known as behavioural economics." □ Improve transition from previous paragraph, defining neoclassical economics should come before saying "these deviations from neoclassical economics" – or should be incorporated into that first sentence
 - Paragraph rewritten to state the neoclassical assumption first ("stable preferences and well-informed decisions in self-interest") before introducing behavioural economics as a deviation from it. [Section I (Introduction), paragraph 2]

- I would suggest incorporating more evidence or examples into first paragraph of intro + Introduction is under-cited
 - Added supporting citations (He et al., 2018; Liu et al., 2018) to the literature-gap paragraph, and lightly expanded the opening health/economic framing. [Section I (Introduction), paragraphs 1 and 6]
- Unclear: “There exists a growing body of literature on both air pollution and behavioural economics individually, but a critical gap persists at their intersection. Where behavioural approaches have been recently applied to environmental risk perception, they tend to examine cognitive biases within a narrow empirical scope, rather than developing an integrated account from various contexts of how the general bias impacts air quality perception as a whole. The systemic application of heuristics—namely the halo effect confirmation bias, and loss aversion—to air quality perception and pollution-related policy design therefore remains underdeveloped in existing literature.”
 - Sentence restructured for clarity and the stray character removed. [Section I (Introduction), literature-gap paragraph]
- The last 6 paragraphs of the intro which outlines the structure of the paper is good but long and slightly repetitive so it should be condensed
 - Condense last 6 paragraphs into a single roadmap paragraph. [end of Section I (Introduction)]
- Consider including “narrative review” in title or abstract so it is clear earlier on the structure that it will take and why there is a methods section
 - Added "Through a narrative review of..." to the Abstract, and the Methodology explicitly states the review "is thus interpretive instead of systematic." [Abstract; Section II (Methodology), paragraph 1]
- Not necessary to say “Studies cited later on in this paper will focus on levels of PM2.5 as a measure of air quality due to the seriousness of its health consequences and geographical ubiquity (Henning, 2024).” – this is talked about in paragraph after and can speak for itself
 - Retained but repositioned directly beside the pollutant discussion so it is not duplicated later, and a supporting sentence on pollutant-specific perception was added. – Section III (Background), paragraph 2]
- Font and size changes throughout, particularly in Loss Aversion section
 - Formatting standardised throughout the section and manuscript.
- Citations are often placed outside punctuation (i.e., . (Rankin, 1969))
 - Citation placement standardised so citations precede the full stop throughout (e.g., "...less polluted than others (Rankin, 1969)."). [throughout, e.g. Section IV (Halo Effect)]

- “I suggest that governments can provide subsidies or tax rebates to real estate projects” – first use of first person, either introduce throughout or change
 - Rewritten in third person: “A proposed application is that governments can provide subsidies...” [Section IV (Halo Effect), real-estate subsidy paragraph]
- Informal wording here:
 - “We could even go as far to say that the success of mitigation policies hinge on thorough understanding...”
 - “Since people love to gain social capital and follow social trends, being eco-conscious will come to be seen as the “in” thing, encouraging widespread adoption of green behaviours through social proof.”
 - Softened to “We might even argue...”, and “the ‘in’ thing” changed to “fashionable,” with a supporting citation (Weiler et al., 2022) added for the social-capital claim. [Section IV (Halo Effect)]
- Be cautious of overstating causality: “The confirmation bias is highly relevant in the context of air pollution, lying in the fact that air pollution has been a recognised problem for millennia”
 - The historical-recognition claim was separated from the confirmation-bias claim into two distinct sentences. [Section V (Confirmation Bias), opening paragraph]
- “Up till this point” – should be until
 - Corrected to “until” [Section VII (Discussion), paragraph 1]
- “tap on this heuristic” should be “tap into this heuristic”
 - Reworded entirely to “policymakers may be able to use this heuristic to increase the effectiveness of pro-environmental policy.” [Section VI (Loss Aversion)]
- Wording is odd here: “These examples serve to nudge individuals to more willingly adopt greener lifestyle habits.”
 - Reworded to “These examples may help in nudging individuals to adopt greener lifestyle habits more readily.” [Section VI (Loss Aversion), closing line]
- “The paper provides conclusive evidence for loss aversion in the environmental context.” – too strong, these findings provide evidence consistent with loss aversion influencing environmental decisions, but other psychological and contextual factors may also contribute
 - Softened to “The paper thus provides evidence for loss aversion in the environmental context.” [Section VI (Loss Aversion), Bartczak paragraph]

Reviewer 2's comments

Title, Keywords, and Abstract:

- Title: The title is engaging and accurately signals the interdisciplinary focus. Consider whether “Analysing” and “Behavioral” should follow one spelling convention throughout the manuscript, since British and American English are currently mixed.
 - British spelling is now applied consistently throughout (Analysing, Behavioural, organisation, judgement, neighbourhood, maximise, etc.), and keyword capitalisation was aligned.
- Keywords: The keywords are relevant, but “Policy Design” should follow the same capitalization style as the other terms. You may also want to consider whether “risk perception” is central enough to include.
 - "Risk perception" added as an eighth keyword.
- Abstract, opening sentence: The opening establishes significance quickly, but “severe, global issue” is slightly general. Consider whether the sentence could identify air pollution as an environmental health risk rather than relying on broad emphasis.
 - Retained but immediately grounded with quantified premature-deaths and economic-cost figures in the same and following sentences.
- Abstract, sentence beginning “Heuristics”: The phrase “due its subjectivity and complexity” needs a grammatical correction and a clearer explanation of what is subjective. Air pollution itself is measurable, while public perception of it is subjective.
 - Corrected to "due to the subjectivity and complexity of individual discernment based on visual and olfactory cues," clarifying that perception (not the pollution itself) is subjective.
- Abstract, sentence beginning “This paper conducts an analysis”: “Conducts an analysis into” is awkward. The sentence would read more clearly if it stated that the paper examines the role of the three heuristics.
 - Changed to "This paper examines the critical role of three key heuristics..."
- Abstract, halo effect sentence: The claim that the halo effect “causes” residents to overestimate air cleanliness is stronger than the reviewed evidence appears to support. Consider using language that reflects association or contribution unless the cited studies establish causation.
 - Softened to "can be linked to residents systematically overestimating..."
- Abstract, confirmation bias sentence: This is clear, but it presents one direction of the bias only. The paper later discusses both optimistic and pessimistic distortions. Consider whether the abstract should reflect that broader application.

- Revised to cover both directions: "entrenched negative (or positive) perceptions of air pollution in the face of objective environmental improvements (or decline)."
- Abstract, loss aversion sentence: The connection to carbon taxes is interesting, but "policies that trigger individuals' loss aversion" is broad. It may help to specify that the issue concerns how costs and incentives are framed and perceived.
 - Specified to "disproportionately undermines policies that do not directly offset upfront investment costs."
- Abstract, sentence beginning "Through a review": The abstract would be stronger if it identified the paper as a narrative review, since the methodology section makes that distinction important.
 - Added "Through a narrative review of foundational psychological experiments and environmental studies..."
- Abstract, final sentence: "Mitigate and/or strategically harness" is cumbersome. Consider choosing one clear construction and stating that the policies either reduce the effects of these biases or use them to improve policy response.
 - Simplified to "designed to either reduce the effects of these biases or harness them to improve policy response."
- Abstract, overall: The abstract captures the main argument well, but it would benefit from a brief indication that the proposed interaction among the three heuristics is conceptual and remains to be tested empirically.
 - The conceptual status is now made explicit in-text at first mention in the discussion ("not...a definitive model but a plausible account"); the abstract itself was left concise per its length constraints, with the full framing preserved in the body.

Section I (Introduction):

- Introduction, opening paragraph, sentence beginning "Air pollution stands": "Environmental causes" does not seem to be the intended term. Consider whether "environmental threats" or "environmental health risks" better expresses the point.
 - Changed to "environmental threats."
- Introduction, opening paragraph, economic-cost sentence: The sentence is a fragment. It also uses highly emphatic wording. Consider integrating the economic cost into the preceding or following sentence and maintaining a more measured academic tone.
 - Rewritten as a complete, more measured sentence.
- Introduction, opening paragraph, sentence beginning "Yet, societal and individual responses": This is a strong transition from the physical problem to the psychological argument. The phrase "singular threat" may be unnecessary because the paper later acknowledges that pollution risk varies across settings.

- "Singular" removed.
- Introduction, behavioral economics paragraph: The contrast between neoclassical and behavioral economics is useful, but it is presented as more absolute than most economic theory would support. Consider acknowledging that neoclassical models often simplify behavior rather than literally assuming that all actors possess perfect information.
 - Reworded to describe neoclassical economics as assuming "stable preferences and well-informed decisions," rather than an absolute claim of perfect information for all actors
- Introduction, behavioral economics paragraph, final citation: The citation is placed as a separate sentence. It should be attached to the claim it supports.
 - Citation now attached directly to the claim it supports.
- Introduction, sentence beginning "This paper argues": The thesis is clear, but "we must" gives it a persuasive tone. A more neutral formulation may be more appropriate for a review article.
 - Reworded to a more neutral formulation: "...can raise the effectiveness of anti-pollution movements."
- Introduction, paragraph beginning "The imperative for this behavioral approach": The sentence stating that the most harmful pollutants are imperceptible is too broad. Some hazardous pollutants can produce visible or sensory cues, while PM2.5 itself is generally not directly visible. Consider narrowing the claim.
 - Narrowed explicitly to PM2.5, the pollutant on which the paper's cited sources focus, rather than pollutants generally.
- Introduction, same paragraph, phrase "human sense": This should be plural or otherwise clarified because the paragraph later refers to sight, smell, and physical discomfort.
 - Corrected to "human senses."
- Introduction, same paragraph: The examples of visual haze, smell, media reports, and discomfort are useful. Consider distinguishing direct sensory cues from socially supplied information, since they may activate different judgment processes.
 - Added a separate sentence: "Socially supplied information, such as that from media reports and hearsay, can distort judgements as well."
- Introduction, paragraph beginning "Understanding these cognitive processes": This effectively explains the policy relevance. "Instrumental data" is unclear in this context. "Objective measurements" or a similarly specific term may be easier to understand.
 - Changed to "objective measurements."

- Introduction, literature-gap paragraph: The paper identifies a meaningful gap, but the claim that the integrated application “remains underdeveloped” needs stronger support. Consider citing one or two prior reviews or explicitly noting the scope of the search that led to this conclusion.
 - Added citations (He et al., 2018; Liu et al., 2018) evidencing the narrow empirical scope of prior work.
- Introduction, literature-gap paragraph, list of heuristics: A comma is missing between “halo effect” and “confirmation bias.” There is also a stray “z” at the end of the paragraph.
 - Both corrected.
- Introduction, roadmap (spanning sections) paragraphs: The manuscript explains the contents several times. These paragraphs could be reduced substantially. One concise roadmap would be enough and would create more space for argument development.
 - Condensed into a single roadmap paragraph.
- Methodology, roadmap paragraph beginning “The paper proceeds as follows”: The section numbering does not match the manuscript. Methodology is Section II, Background is Section III, and the heuristic sections begin with Section IV. Please correct all section references after the final structure is set.
 - Manuscript renumbered consistently: Introduction (I), Methodology (II), Background (III), Halo Effect (IV), Confirmation Bias (V), Loss Aversion (VI), Discussion (VII), Conclusion (VIII); the roadmap text now matches.
- Methodology, sentence beginning “The discussion of loss aversion”: “Perceptive losses” is not the right term. Consider wording that refers to perceived losses, upfront costs, or lifestyle sacrifices.
 - The sentence in question was part of a repetitive roadmap paragraph that was removed during condensation; the corrected term (“perceived losses”/lifestyle sacrifices) is used where the concept is introduced substantively in Section VI.
- Introduction, overall: The introduction makes a strong case for the topic, but it would benefit from a clearer separation among three elements: the public health problem, the behavioral theory, and the specific contribution of this review. At present, these elements overlap and occasionally repeat one another.
 - Improved through the paragraph-level edits above and the condensed, corrected roadmap, which now states the contribution and structure in one clear pass rather than several overlapping ones.

Section II (Methodology):

- Methodology, opening paragraph: The narrative review design is appropriate for the interdisciplinary aim, and the rationale is explained well. Consider adding a brief statement

that the review is interpretive rather than systematic so readers understand the limits of reproducibility.

- Added "...is thus interpretive instead of systematic."
- Methodology, search strategy: The databases and organizations are listed clearly. The methodology would be stronger if the author reported when the searches were conducted and whether the same search combinations were used across databases.
 - Added: searches "conducted from June 2025 to November 2025" and "the same search combinations were used across databases."
- Methodology, organization name: "World Organisation" appears to be incomplete. It should identify the World Health Organization accurately. Please be sure to not alternate between US and British Spellings. Keep with what the manuscript allows.
 - Corrected to "World Health Organisation (WHO)"; British spelling maintained throughout.
- Methodology, search terms: The listed terms cover the core concepts, but policy-related terms are absent even though policy design is central to the title and conclusion. Consider whether the search strategy included terms such as environmental policy, risk communication, incentives, or behavioral intervention.
 - Added "environmental policy" to the list of search terms.
- Background, sentence beginning "No strict date restriction": The rationale for including foundational work is sound. The wording around "Kahneman and Tversky's on prospect theory" and "Wason's on the rule discovery task" is incomplete and needs grammatical correction.
 - Corrected to "Kahneman and Tversky's paper on prospect theory (1979) and Wason's paper on the rule discovery task (1960)."
- Background, inclusion criteria: The criteria are reasonable but remain broad. Consider clarifying how credibility was judged for institutional reports and how methodological transparency was assessed.
 - Inclusion criteria tightened to specify publication in peer-reviewed journals or by "globally recognised international institutions"; a fuller credibility-assessment protocol was judged beyond the scope of a narrative review and is flagged as a direction for more systematic future work.
- Background, exclusion criteria: "Purely anecdotal or non-empirical commentary" is clear, but the manuscript later makes policy proposals that are not directly tested. It would help to distinguish literature-selection criteria from the author's own conceptual recommendations.

- Added closing sentence: "Simple policy recommendations were then constructed to address the identified gaps directly," separating the review's selection criteria from the author's original proposals.
- Background, thematic organization: This is one of the strongest parts of the methodology because it explains how foundational studies and environmental applications were linked. Consider making the same sequence visible in each substantive section through consistent subheadings.
 - Worked to make organising principle more consistently legible across Sections IV–VI, though chose to reinforce it through structure and language rather than inserting formal subheadings. Each section now opens with the foundational psychological study, transitions explicitly into the air-quality-specific application, and closes with policy implications clearly flagged as proposals – for example, the Halo Effect section moves from Nisbett and Wilson's experiment, to Boso et al. (2020) and Hofflinger (2019) on air quality specifically, to the real-estate subsidy and Green-Status proposals, with transition sentences marking each shift ("Previous research has examined the impact of the halo effect on individual perception of air pollution in specific contexts...").
- Background, policy implications sentence: "Derived inductively" is a meaningful methodological claim. Consider explaining how the author moved from the evidence to the proposed recommendations, especially where a policy idea is original rather than directly supported by a reviewed intervention study.
 - Treated this as two separate steps and said so explicitly in the Methodology (policy implications are drawn from "patterns observed across studies within each thematic cluster," and recommendations are then "constructed to address the identified gaps directly"). The added clause flags that the recommendations are my own construction, not something read directly off the cited studies. I carried this distinction into the sections themselves rather than leaving it abstract: each proposal in Sections IV–VI is now explicitly labelled as a proposal, and where it goes beyond what any single reviewed study tested, I have added a caveat to that effect (e.g., the equity caveats on the real-estate subsidy and Green-Status badge proposals, and the distributional caveats on the green-bonus and streak-bonus proposals).
- Methodology, overall: The section is a good start, but the review would be easier to evaluate if it included the approximate number of sources screened and retained, even if a formal systematic-review flow diagram is not required.
 - Not incorporated in this revision, as the narrative-review design (per the point above) does not use a systematic screening protocol; the search databases, date range, and terms are now reported in full to maximise the transparency that is available within this design.

Section III (Background):

- Background, opening paragraph: This paragraph clearly establishes the purpose of the section. “Personalised human interpretation of objective reality” is somewhat abstract. Consider using more direct wording about how measured pollution levels are interpreted by individuals.
 - Reworded to "transforming measured pollution levels into varying interpretations from individual to individual."
- Background, definition of air pollution: The definition is appropriate. Consider whether quotation marks and citation style are consistent with the rest of the manuscript.
 - Page number added: "(WHO, 2024, p. 1)."
- Background/Methodology, pollutant overview: This is useful background, but the paper later focuses mainly on PM2.5. Consider explaining earlier why PM2.5 is the primary example while still acknowledging that different pollutants are perceived differently.
 - The PM2.5-focus sentence now appears immediately alongside the pollutant list, and a new sentence was added: "Perception of air pollution varies from pollutant to pollutant given their different physical characteristics."
- Background, global burden paragraph: The health statistics establish significance, but several figures are presented in rapid succession. Consider grouping them by global burden, regional burden, and disease outcomes so the reader can follow the evidence more easily.
 - Explicit transition markers added ("Regionally..."; "Additionally, disease outcomes are severe...") to separate the three categories.
- Background, sentence beginning “As a whole, air pollution accounts”: “The culprit of” is informal and grammatically awkward. The relationship between the 1-in-8 estimate and the 7.9 million figure should also be checked for consistency and source year.
 - "The culprit of" removed; a "(2025)" timestamp added to the 1-in-8 estimate for consistency with the 7.9 million (2023) figure.
- Background, sentence beginning “In particular, studies have shown”: The list of diseases is broad. Consider distinguishing conditions with well-established causal evidence from those for which the association is still emerging.
 - Revised to distinguish diseases with more consistently demonstrated increased risk (“associated with the increased risk of cardiovascular and respiratory diseases, lung cancer, stroke, pneumonia, depression, and diabetes”) from others with emerging associations (“also been linked to several other conditions, such as asthma, bronchitis, breast cancer, etc., but concrete causal linkage warrants further investigation”).

- Background, economic-cost paragraph: The range of international examples adds breadth, but the paragraph is overloaded. Consider separating global estimates from country examples and limiting the number of statistics to those most relevant to the argument.
 - Global and country-level figures are now marked with explicit transitions ("Globally, estimates project..."; individual country sentences retained separately for UK/India/Southeast Asia).
- Background, India sentence beginning "Had the country managed": The grammar is incorrect, and the counterfactual claim needs especially careful wording. Check the original source to make sure the time period and estimated GDP effect are represented accurately.
 - The unverifiable counterfactual claim was removed; the India paragraph now reports only the sourced \$95 billion/3% GDP figures.
- Background, sentence beginning "In Southeast Asia": "Snowballing to a hefty" is journalistic. A more neutral expression would fit the rest of the review.
 - Changed to "with costs amounting to 600 billion USD."
- Background, closing sentence "Undoubtedly": Avoid absolute or rhetorical language. The evidence supports a strong economic connection without needing this qualifier.
 - Sentence removed.
- Background, paragraph beginning "While the physical and sociological facts": "Sociological facts" is not clearly defined. Consider whether "health and economic evidence" better reflects the preceding discussion.
 - Changed to "the physical, health and economic evidence of air pollution."
- Background, PM2.5 description: The sentence is grammatically difficult and could mislead readers into thinking PM2.5 is always invisible in every context. Consider clarifying that individual particles are not visible and that ambient haze is an imperfect proxy.
 - Reworded to specify that "individual particles" of PM2.5 are invisible and that ambient haze is an imperfect proxy, rather than implying the pollutant itself is always undetectable.
- Background, sentence ending with the Fang citation: The citation is separated from the sentence by a period. Please standardize citation placement throughout.
 - Citation moved inside the sentence it supports.
- Background/Halo Effect, System 1 and System 2 paragraph: The framework is relevant, but the explanation is repetitive and includes a stray superscript-style "15." The sentence beginning "This operation" does not have a clear referent.

- Stray superscript numeral removed and the ambiguous "this operation" referent eliminated through rewording.
- Halo Effect, sentence beginning "Heuristics are the rules of thumb": This is clear, but "generally" is repeated. Consider revising for smoother flow.
 - Changed to "...they are typically effective but come with a trade-off."
- Halo Effect, cue-substitution paragraph: The examples are useful and directly support the paper's thesis. The question mark in "Does it smell bad?" should not be followed by a period.
 - Punctuation corrected.
- Halo Effect, same paragraph: The claim that physical cues create a shift from "data-rich" to "cue-driven" decision making is plausible but would benefit from a direct citation.
 - Citation added (Boso et al., 2025).
- Halo Effect, final sentence: "I further probe into" is informal and shifts into first person. Consider a more neutral transition.
 - Changed to "The following three sections look into three specific cognitive biases..."
- Background, overall: The section successfully establishes why air pollution is difficult to perceive. It would be stronger if the health and economic statistics were condensed and more space were given to the specific psychological problem of judging invisible or delayed risks.
 - Addressed in part through the trimming above (removal of the "Undoubtedly" sentence, the India counterfactual, and the "sociological facts" wording); a full condensation of the statistics was not undertaken in this revision, as several figures were specifically requested elsewhere in the review (e.g., regional grouping, source-year consistency) and removing them further risked contradicting those points.

Section IV (Halo Effect):

- Halo Effect, opening example: The everyday example is accessible, but the second-person framing is conversational. Consider introducing the concept in a more academic voice while keeping the example brief.
 - Changed from direct address ("If you were to meet...") to a more academic third-person framing ("If one were to encounter another individual...").
- Halo Effect, sentence beginning "This is due to the human tendency": "Base our general judgement off" is informal. The statement also implies that one impression always determines all later judgments. Consider more cautious wording.

- Reworded to "form an overall judgement of a person or object from a single lasting impression," removing the absolute implication.
- Halo Effect, sentence beginning "The halo effect was first coined": The effect was named or described by Thorndike, not "first coined by" the person. Check the historical phrasing.
 - Corrected to "first described by."
- Halo Effect, quoted definition: Confirm the page number and citation style. The manuscript alternates among "Pg," parenthetical author-date citations, and numbered references.
 - Standardised to "(1977, p. 1)" format, consistent with citation style used throughout.
- Halo Effect, Nisbett and Wilson experiment: The study is explained clearly, but the amount of procedural detail may be more than the review needs. Consider reducing the description and using the saved space to explain the limitations of applying a person-perception experiment to place-based air quality judgments.
 - Description condensed, and an explicit limitation added noting the study "does not draw a direct link to place-based air-quality judgements."
- Halo Effect, figure caption: Standardize figure-caption formatting and verify permission or fair-use requirements for reproducing the figure. The image is low resolution and may not reproduce well in publication.
 - Caption formatting standardised (citation style, "p." notation); image removed.
- Halo Effect, interpretation paragraph: Verb tense shifts from "rated" to "rate." Please keep the description in past tense.
 - Corrected to consistent past tense.
- Halo Effect, sentence beginning "The results quantitatively prove": "Prove" is too strong for a single experiment. Consider stating that the results support or demonstrate evidence of the halo effect under the study conditions.
 - Changed to "quantitatively supporting the existence of a halo effect in play."
- Halo Effect, paragraph beginning "Fundamentally": This makes the transition to air quality effectively. The phrase "global evaluations are powerful enough to trump other objective data" is stronger than the cited evidence supports and should be qualified.
 - Softened to "can be powerful enough to override other available objective data."
- Halo Effect, sentence beginning "As a result, vigorous public awareness campaigns": The sentence makes a large causal leap from the halo effect to persistent mortality. The reviewed evidence does not establish that the halo effect explains mortality trends. This point needs substantial caution or removal.

- Rewritten to remove the mortality-rate claim; now reads that awareness campaigns and research "have yet to inspire a systemic overhaul of pollutive practices," without asserting a causal link to mortality trends.
- Halo Effect, sentence beginning "This highlights firsthand perception": Calling perception "the root of the issue" overlooks structural sources of pollution, regulation, housing, income, and access to alternatives. Consider framing perception as one contributing barrier rather than the root cause.
 - The "root of the issue" framing was removed from the text.
- Halo Effect, sentence beginning "Previous research has examined the effect of the halo effect": Avoid repeating "effect."
 - Reworded to "Previous research has examined the impact of the halo effect on individual perception..."
- Halo Effect, Chile study paragraph: "As cited" is unclear, and the claim about wood-burning stoves across the Global South is very broad. Consider limiting the statement to the location and source being discussed.
 - Claim narrowed to Chile specifically, removing the "across the Global South" generalisation.
- Halo Effect, same paragraph: Persistent pollution despite intervention may suggest several explanations, not only a halo effect. Fuel availability, cost, cultural practice, housing quality, and enforcement may also matter. The interpretation should acknowledge alternative explanations.
 - Added "Although, alternative explanations such as fuel availability cost, culture practice, and enforcement, are to be acknowledged as well."
- Halo Effect, questionnaire description: Several phrases are incomplete, including "a short five-item evaluating." Consider summarizing the instrument more concisely and clearly.
 - Questionnaire description rewritten in full, clearly summarising the three-part structure (brief rating, extended rating, and demographic/health questions).
- Halo Effect, PM2.5 paragraph: "High levels of particulate matter levels" is repetitive. Units should be formatted consistently as $\mu\text{g}/\text{m}^3$, and the WHO guideline should be checked against the relevant averaging period.
 - Reworded, units standardised to $\mu\text{g}/\text{m}^3$, and a healthy/unhealthy classification against the WHO threshold added along with daily-mean values (33.07 and 35.37 $\mu\text{g}/\text{m}^3$) alongside the reported peaks.
- Page 10, home halo calculation: The explanation of 80%, 70%, and the resulting 10% halo effect is difficult to interpret. Consider clarifying what each percentage represents and whether the difference is statistically significant.

- Expanded with the underlying rating values (home 6.08, neighbourhood 4.33, city 3.83 on a 1–9 scale) and a significance result ($p = 0.001$) so the percentages' basis is clear.
- Page 10, Hofflinger model paragraph: The description provides useful context, but it is lengthy. Consider naming the model type and reporting only the predictors that were significant or central to the interpretation.
 - Condensed; reporting focused on the central finding (61% vs. 7% home-vs-city comparison and the positive mean halo-effect variable). The specific model type was not additionally named in this revision.
- Page 10, neighborhood halo sentence: The sentence is grammatically difficult and the Rankin citation is placed after a period. Clarify the relationship among the home halo, neighborhood halo, and localization of risk.
 - Sentence rewritten, citation moved inside the sentence, and a synthesis sentence added: "Both studies support the finding that people rate environmental conditions in their immediate surroundings as significantly better than broader or more distant areas."
- Page 10, policy transition beginning "Since the significance": "Steadily supported" is vague. Consider stating more specifically what the evidence supports and what remains uncertain.
 - Changed to "well established through quantitative studies."
- Page 10, sentence beginning "We could even go as far": This is conversational and makes a broad claim. Policy effectiveness depends on more than public attitudes alone.
 - Softened to "We might even argue that the success of mitigation policies is largely dependent on..."
- Halo Effect, real-estate subsidy proposal: This is creative, but the recommendation is not directly tested by the studies reviewed. Consider labeling it clearly as a proposed application and discussing possible equity concerns, including whether subsidies for fashionable neighborhoods could widen existing environmental inequalities.
 - Explicitly labelled as "a proposed application," with an added equity caveat noting that fashionable neighbourhoods are typically more affluent and that subsidy provision risks neglecting less well-off areas.
- Page 11, Green-Status badge proposal: The idea is understandable, but it needs stronger grounding in evidence on social norms, digital badges, or public recognition. Consider whether privacy, verification, unequal platform access, and performative behavior should be acknowledged.
 - Added supporting citations (Casillia and Hickey, 2016; Weiler et al., 2022) and a caveat acknowledging "issues of performative behaviour and unequal platform access." Privacy and verification specifically were not separately elaborated.

- Page 11, phrase "people love to gain social capital" and "the in thing": These expressions are too informal for an academic review and rely on generalization.
 - Changed to "individuals are inclined towards accumulating social capital (Weiler et al., 2022) and follow social trends, being eco-conscious will come to be seen as fashionable."
- Halo Effect section, overall: The section has a clear structure and a strong environmental application. Its main weakness is that the policy claims sometimes move too quickly from association to intervention. A clearer distinction among evidence, interpretation, and proposal would improve the section considerably.
 - Addressed through the individual edits above, which consistently label each policy idea as a proposal and separate it from the evidence it draws on.

Section V (Confirmation Bias):

- Page 11, opening paragraph: The concept is introduced accessibly, but the sentence is long and contains a parallel-structure error between "tendency to cherry-pick" and "rejecting." Consider a more concise academic definition before the quotation.
 - Tightened into a more concise academic definition ahead of the Nickerson quotation.
- Page 11, Nickerson quotation: The comma placement around the author attribution is incorrect. Confirm the page reference and quotation formatting.
 - Standardised to "(1998, p. 1)" with corrected punctuation.
- Confirmation Bias, Wason task: The explanation is generally clear, but it could be shortened. The key point is that participants sought confirming cases rather than tests that could falsify their rule.
 - Description tightened to focus on participants' tendency "to seek supporting rather than disconfirming evidence."
- Page 11, sentence beginning "Each round": "Subsequently feedback to whether" is grammatically incorrect.
 - Corrected to "...then asked to infer the rule by proposing their own triples and receiving feedback after each attempt."
- Page 11, sentence beginning "The methods participants employed": "Bore affirmative evidence" is awkward. The interpretation would be clearer if it described how confirmatory testing preserved an incorrect hypothesis.
 - Reworded to "Their tendency to seek supporting rather than disconfirming evidence exhibits use of a confirmation bias to arrive at their eventual conclusions."

- Confirmation Bias, Lord study: The study is useful for showing belief polarization. Verify the author names and publication year, since the bibliography lists Lord, Ross, and Lepper (1979), while the text states Lord, Lee, and Lepper (1981).
 - Corrected throughout the text to "Lord, Ross, and Lepper (1979)," matching the bibliography.
- Page 12, percentage changes: Explain what the percentages represent and whether they are percentage point changes or a transformed attitude measure. Without this context, the numbers may be misleading.
 - Added "The following results are presented as percentage point changes."
- Page 12, sentence beginning "Such data strongly supports": Subject-verb agreement should be corrected. "Will lead" is also too definite for one study.
 - Corrected to "Such data strongly supports the view that mixed data can lead to increased polarisation rather than uncertainty and moderation."
- Page 12, paragraph beginning "The confirmation bias is highly relevant": The logic is not yet clear. The fact that pollution has been recognized for a long time does not itself establish confirmation bias. Consider beginning with the persistence of prior beliefs despite updated evidence.
 - The historical-recognition claim was separated into its own sentence, distinct from the confirmation-bias argument that follows.
- Page 12, citation spelling: "Shwartz" and "Schwartz" are both used. Standardize the author's name and verify the publication year, which also differs between the text and bibliography.
 - Standardised to "Schwartz" and corrected to (2004) in both the in-text citations and the bibliography.
- Page 12, sentence beginning "Previous research, albeit sparse": "Substantiates the correlation" may be too strong given the observational nature of the cited studies. Consider saying that the research is consistent with or suggestive of confirmation bias.
 - Softened to "is consistent in supporting the correlation between confirmation bias and air pollution."
- Page 12, U.S. air-quality paragraph: The comparison is interesting, but the data appear historical. Consider adding a more recent source or explaining why the older perception data remain relevant to current policy design.
 - Added "Although the above data is historical, the long period over which the study was conducted offers insight into the enduring nature of confirmation bias, which is relevant to current policy design." A more recent source was not additionally added.
- Page 12, sentence beginning "Ironically enough": Remove conversational commentary and present the contrast directly.

- Removed and replaced with a direct contrast ("In fact, actual ozone level data reflect...").
- Page 12, claim that most American cities enjoy relatively good air quality: This statement is broad and may obscure local disparities, episodic exposure, and pollutants other than ozone. It needs qualification.
 - Qualified to "relatively better air quality than in the past," rather than an unqualified claim of good air quality.
- Page 12, sentence beginning "It is likely the presence of confirmation bias": This is an inference, not a direct finding. Clearly label it as the author's interpretation and acknowledge other explanations, such as media exposure, uneven local improvement, or distrust of official data.
 - Reframed as "Based on the evidence presented, it is inferentially likely...", with an added sentence acknowledging media exposure, uneven local improvement, and distrust of official data as alternative explanations.
- Confirmation Bias, King study: The socioeconomic and racial findings are important but sensitive. Take care not to imply that racial composition itself is a simple cause of inaccurate perception. The discussion should distinguish measured neighborhood characteristics, structural inequality, place stigma, and individual bias.
 - Statistics were moved into a neutrally-framed table (Table 1), "racial/ethnic phenotypes" was replaced with "individual bias of racial and ethnic composition," and an added sentence acknowledges alternative drivers such as news coverage patterns and literacy levels.
- Page 13, long sentence beginning "Key results are as follows": The sentence contains several statistics and should be broken into shorter units or presented in a table.
 - Converted into Table 1 ("Summary of statistics for socioeconomic groups by proportion by air quality rating").
- Page 13, sentence beginning "Thus, social cues": "Emerge as the key predictors" is too definitive unless the statistical model establishes this after adjustment for objective pollution and other covariates. Consider reporting the study's actual conclusion more carefully.
 - Softened to "emerge as a likely key predictor," paired with an explicit correlational-data caveat.
- Page 13, "racial/ethnic phenotypes": This terminology is not appropriate for neighborhood demographic composition. Consider using racial and ethnic composition or another term consistent with the source.
 - Changed to "individual bias of racial and ethnic composition."

- Page 13, policy transition beginning "Fully grasping": "Leverage on" should be corrected. More importantly, the ethics of using confirmation bias should be discussed, particularly the difference between value-aligned communication and manipulative targeting.
 - Grammar corrected, and a new closing paragraph added distinguishing ethical value-aligned communication from manipulative targeting, requiring messaging to remain "factually correct and scientifically consistent."
- Confirmation Bias, conservative and progressive messaging examples: The examples are plausible, but they risk oversimplifying political groups and reinforcing stereotypes. Consider grounding the recommendation in audience-segmentation or framing research and acknowledging variation within groups.
 - Grounded with citations (Hurst and Stern, 2020; Boso et al., 2025; Ditto et al., 2017; Washburn and Skitka, 2017; Ballew et al., 2019), and an explicit caveat added: "The above suggestions rely on a simplification of political groups and acknowledge variation within groups."
- Page 14, trusted-messenger proposal: This is one of the more practical recommendations. It would be stronger with evidence that trusted messengers improve environmental-health communication or behavior.
 - Supported with a citation (Shen et al., 2023) on trusted community-based organisations in health communication.
- Confirmation Bias section, overall: The section presents strong foundational evidence and a useful policy application. It needs more caution when using observational studies to infer a specific cognitive mechanism, and the ethical limits of targeted messaging should be addressed.
 - Addressed cumulatively through the caution/caveat edits above and the new ethics paragraph on targeted messaging.

Section VI (Loss Aversion):

- Page 14, opening coin-toss example: The example is memorable, but direct address and rhetorical questions create a conversational tone. Consider summarizing the demonstration in a more formal way.
 - Reframed in the third person around Kahneman's classroom demonstration, removing direct reader address.
- Page 14, sentence beginning "From this demonstration": The claim that people place "twice as much value" on losses should be presented cautiously. The approximate loss-aversion coefficient varies by context and is not a universal constant.
 - Softened to "suggesting a tendency to place more value on losing money as opposed to gaining money," without asserting a fixed 2:1 ratio.

- Page 14, prospect theory paragraph: The theoretical explanation is relevant, but the distinction between loss aversion and risk aversion should be made especially clear. These are related but not interchangeable concepts.
 - Added explicit sentence: "It is important to note the distinction between loss and risk aversion, where the latter refers specifically to a preference for the certain outcome over a gamble with the same expected value (Menezes and Hanson, 1970)."
- Page 14, quoted definition: The quotation appears to use "gain" where "gains" may be expected. Verify the wording against the original source.
 - Corrected to "gains," matching the original source.
- Loss Aversion, Karle experiment: The experiment is explained in considerable detail. Consider condensing the lottery mechanics and focusing more directly on what the design shows about risk preference versus loss aversion.
 - Description tightened.
- Page 14, citation year: The text gives Karle et al. (2013), while the bibliography lists 2015. Verify and standardize.
 - Standardised to 2015 throughout the text and bibliography.
- Page 15, interpretation of Series A: The description of risk aversion appears reversed. Choosing a sure amount below the lottery's expected value may indicate greater risk aversion, but the classification needs careful checking against the original study.
 - Corrected; choosing a sure payment below the lottery's expected value is now correctly identified as risk-averse behaviour, and choosing the lottery over sure payments above the expected value as risk-seeking; the reported figure (23.3% risk-seekers) now matches this corrected classification.
- Page 15, Series B values: The list says there are six choices but provides seven values of R. Check the original experiment and correct either the number of choices or the list.
 - Corrected to "seven choices," matching the seven listed values of R.
- Page 15, sentence beginning "An individual that, additionally": This is difficult to follow and should be simplified.
 - Simplified to "an individual that exhibits loss aversion, and thus risk-aversion, would be willing to accept a payment of zero more often for any given R."
- Page 15, environmental-action paragraph: The connection between immediate private costs and future environmental benefits is useful. However, the statement that people choose efficient products only when payback is very short is broad and should reflect the specific scope of the cited review.

- Presented as a direct, page-referenced quotation from Kollmuss and Agyeman (2002, p. 249) rather than an unqualified paraphrase.
- Page 15, sentence stating people value the loss of time more than environmental benefit: This interpretation may conflate time preference, uncertainty, liquidity constraints, habits, and loss aversion. Consider acknowledging competing explanations.
 - Added explicit limitation: "A limitation of this analysis is that it does not disentangle loss aversion from other alternative explanations such as heterogeneity in time preferences and liquidity constraints."
- Loss Aversion, Bartczak study: The study is relevant, but the layout appears broken across lines and needs formatting cleanup. The treatment descriptions can also be condensed.
 - Treatment descriptions tightened and layout cleaned up.
- Page 16, sentence beginning "The paper provides conclusive evidence": Avoid "conclusive." A field experiment can provide evidence consistent with loss aversion, but it does not settle the broader question across all environmental contexts.
 - Changed to "The paper thus provides evidence for loss aversion in the environmental context."
- Page 16, Knobloch study description: Calling the study observational is inaccurate if it is a simulation or modeling study. Use the study design described by the authors.
 - Corrected to "a simulation study."
- Page 16, claim that cultural factors do not predominantly influence heating decisions: This needs support and may be difficult to defend across 59 regions.
 - Supported with a citation (Michelsen and Madlener, 2013).
- Page 16, carbon tax and subsidy paragraph: The explanation of taxes as the "gain element" and subsidies as the "loss element" appears conceptually confused. A carbon tax is generally experienced as a cost, while a subsidy reduces an upfront cost. Recheck the model's framing and reference points.
 - The confusing "gain element"/"loss element" labels were removed; the text now simply explains that subsidies are more effective because they directly offset the upfront-cost pain point for loss-averse consumers, while taxes do not.
- Loss Aversion, tax comparison: The quantitative comparison is useful but should be reported with the model assumptions and scenario context. Avoid generalizing the result beyond the simulation.
 - Framed explicitly as a model projection ("projections show that a carbon tax of 200 €/tCO₂...").

- Page 17, phrase "double the tax": This is too informal and may overstate what the model establishes. State the scenario comparison precisely.
 - Reworded to "a tax rate about double that of ideal projections (without loss aversion) is required to attain similar outcomes."
- Page 17, sentence beginning "Having proven": The section has not proven that loss aversion governs all pollution-related behavior. Use more cautious language.
 - Changed to "Taken together, these findings suggest that loss aversion may play an important role in environmental pollution and inaction."
- Page 17, health-insurance analogy: The cross-domain example is creative, but transfer from health incentives to household emissions should not be assumed. Consider explaining why the mechanism is expected to generalize and what practical differences may limit it.
 - Added explanation of why the mechanism is expected to generalise: "the considerations underpinning pro-health and pro-environmental behaviour are similar. Both rely on long-term commitment for delayed gains..." Practical limiting differences were not separately elaborated.
- Page 17, green-bonus proposal: The recommendation needs discussion of measurement, household size, income, baseline emissions, renter control over appliances, and the risk of penalizing households with limited ability to change consumption.
 - Expanded with an explicit distributional-risk discussion covering household size, income-driven "regressive outcome" risk, and renters' limited control over efficiency measures, plus a call for adjusted thresholds. Measurement methodology and baseline-emissions specifics were not further detailed.
- Page 17, commuter streak proposal: Capitalization of "Commuters" is unnecessary. The example should also consider accessibility, disability, rural transit availability, and whether losing a fare discount could disproportionately affect lower-income commuters.
 - Capitalisation corrected, and expanded to address fixed-quota assumptions, disability and rural/low-density transit constraints, and the disproportionate impact of losing the discount on lower-income commuters.
- Loss Aversion section, overall: This section has a strong policy focus, but it needs the most technical checking. The manuscript should verify the experimental classifications, distinguish loss aversion from risk aversion and time preference, and reassess how taxes and subsidies are described within the reference-point framework.
 - Addressed via the specific technical corrections above: the Series A/B classification fixes, the explicit loss-aversion/risk-aversion distinction, the time-preference/liquidity-constraint limitation, and the removal of the confused tax/subsidy "gain"/"loss" labelling.

Section VII (Discussion):

- Page 18, section numbering: This section is labeled V even though earlier sections already use V and VI. Renumber the entire manuscript after the final structure is confirmed.
 - Entire manuscript renumbered I–VIII; Discussion is now correctly Section VII.
- Page 18, opening paragraph: The integrative aim is the most original part of the manuscript. The sentence beginning "This section proposes" is grammatically incomplete and should be corrected.
 - Corrected to "This section proposes a potential account of how such an interaction might unfold."
- Page 18, claim of a "perpetual compounding effect": This is stronger than the evidence supports. The reviewed studies do not directly test the joint operation of all three heuristics. Continue framing the model as a hypothesis or conceptual proposition.
 - Softened to "a possible compounding effect," explicitly framed as "a mutually reinforcing relationship not yet explored" rather than an established finding.
- Page 18, sequential-cycle paragraph: The author appropriately states that the model is not definitive. Consider adding a small conceptual figure showing the proposed sequence and feedback loops, while labeling it clearly as a proposed framework.
 - Added Figure 1 ("Proposed framework") to visually present the sequence as a proposal rather than an established finding.
- Page 18, halo-effect paragraph: The theoretical rationale for placing the halo effect first is understandable, but the sequence may differ by context. Prior beliefs or loss concerns could precede place attachment. Discuss alternative ordering rather than implying one universal pathway.
 - Added "However, alternative sequencing remains possible as prior beliefs or loss concerns can precede place attachment."
- Page 18, phrase "in a more favorable direction that objective data warrants": "That" should be "than."
 - Corrected.
- Page 18, sentence beginning "The critical point": This is a comma splice and should be divided or joined correctly.
 - Corrected.
- Page 18, confirmation-bias paragraph: The examples are clear and help the reader understand the proposed mechanism. Because they are hypothetical examples, make that status explicit.
 - Added "A hypothetical example is that on a clear, sunny day..."

- Discussion, loss-aversion paragraph: The transition is logical, but "functions are the final barrier" contains a grammatical error. "Corrected perception" and "corrected behavior" also sound overly binary; behavior can change partially.
 - Corrected to "functions as the final barrier," and "corrected" changed to "changed" throughout to reduce the binary implication.
- Page 19, synthesis paragraph: "More intractable that" should be "more intractable than." Consider stating what empirical design would be needed to test whether the combined effect is additive, interactive, or sequential.
 - Grammar corrected; a new closing paragraph proposes a factorial design and specific testable predictions for distinguishing these possibilities.
- Page 19, phrase "suggestive, if not conclusive": Since the manuscript later acknowledges that the cases do not prove interaction, "if not conclusive" should be removed.
 - "If not conclusive" removed.
- Page 19, Chile case: The synthesis is thoughtful, but the interpretation attributes persistence to confirmation bias and loss aversion without direct measures of either mechanism. Label these as plausible explanations and retain alternative structural explanations.
 - Added "An important caveat is that their data is purely correlational so alternative structural explanations are plausible."
- Page 19, "Padre Las Casa": Check the place name. The earlier section uses Padre Las Casas.
 - Corrected throughout.
- Page 19, "systemically overestimated": The intended word appears to be "systematically."
 - Corrected.
- Discussion, Chicago case: "Aur quality" is a typographical error. More importantly, the claim that loss aversion discouraged investment is speculative and is not measured in the cited study. This should be presented as a research question rather than an inferred mechanism.
 - Typo corrected; reframed as "An open research question is that loss aversion may further have discouraged investment...", presenting it as a question rather than an inferred mechanism.
- Page 20, policy implications paragraph: The example of belief change without behavior change is useful. Consider linking this distinction to established models such as the intention-behavior gap if relevant literature supports it.
 - Citation added (Faries, 2016) linking this distinction explicitly to the intention-behaviour gap literature.

- Page 20, final paragraph: The last sentence repeats "potentially complementary component of a broader behavioral strategy" and appears unfinished. The proposed sequence also stops after confirmation bias and does not complete the loss-aversion component.
 - Final paragraphs rewritten and extended, completing the loss-aversion component of the sequence and adding a new paragraph proposing a factorial experimental design with specific testable predictions.
- Discussion, overall: This section provides the manuscript's clearest original contribution. It should be expanded as a conceptual framework but separated carefully from the empirical findings. A short subsection on testable propositions, alternative pathways, and research design would make the contribution much more convincing.
 - Addressed through the additions above: the explicit conceptual/hypothesis framing, the alternative-ordering caveat, the correlational caveats on both case studies, and the new closing paragraph on a factorial design with testable predictions.

Section VIII (Conclusion):

- Page 20, opening sentence: "Compelling evidence" and "profoundly shape" are stronger than warranted across all three domains. Consider stating that the reviewed evidence indicates or suggests meaningful influence.
 - Softened to "has suggested meaningful influence."
- Page 20, punctuation around heuristic list: The manuscript uses dash characters inconsistently. Replace them with commas or parentheses and standardize punctuation throughout.
 - Standardised throughout the Conclusion (and manuscript).
- Page 20, sentence beginning "Findings explored": The wording is slightly rhetorical. Consider summarizing the practical implication directly.
 - Retained with tighter framing of the practical implication in the surrounding sentences.
- Conclusion, heuristic-summary paragraph: This is a useful recap, but several claims remain causal. For example, confirmation bias may contribute to persistent perceptions, but the observational studies do not isolate it experimentally.
 - Softened throughout ("may help to explain"; "offers a plausible mechanism"), with an explicit caveat added that confirmation bias "cannot be isolated as the cause of these persistent perceptions" given the observational data.
- Page 21, phrase "less than proportionate social impacts": Clarify what kind of impact is meant, such as public recognition, trust, behavior, or policy support.
 - Specified as "generate less policy-support response than the scale of improvement would warrant."

- Page 21, interaction paragraph: "Prevents updates perception" should be corrected. "Casual orderings" should be "causal orderings."
 - Corrected to "prevents updated perception" and "causal orderings."
- Page 21, future-research sentence: This is a strong and necessary limitation. Consider moving part of this caution earlier into the Discussion so the speculative model remains appropriately framed throughout.
 - The factorial-design and testable-predictions discussion was added directly to the Discussion (Section VII) rather than appearing only in the Conclusion, so the model is framed as speculative before the reader reaches the Conclusion.
- Page 21, policy paragraph: "Effective countervailing policies work with, not against" reads as a settled conclusion. Consider saying that policies may be more effective when they account for these patterns.
 - Softened to "Countervailing policies may be more effective when they work with, not against these ingrained psychological patterns."
- Page 21, sentence beginning "As proposed, this can involve": The transition is awkward. The three proposals could be presented as a concise list or parallel sentences.
 - Restructured into an explicit "First,... Second,... Third,..." sequence.
- Page 21, loss-framed incentives: The example of forfeited bonds raises ethical and equity questions. Consider acknowledging that loss-framed designs may burden vulnerable groups or reduce trust if they are perceived as punitive.
 - Added such designs "may disproportionately burden lower-income households for whom the forfeited amount represents a greater relative cost, and may erode public trust if perceived as punitive rather than supportive."
- Conclusion, final paragraph: "Beget real, effective action" is overly rhetorical. The conclusion would be stronger if it ended with a precise statement of what this review contributes and what evidence is still needed.
 - Final paragraph rewritten to end with a precise statement of the review's contribution and the specific empirical gap that remains, including whether interventions can be implemented "without disproportionately burdening populations least able to absorb their costs."
- Conclusion, overall: The conclusion accurately reflects the paper's main themes, but it should distinguish more carefully among established evidence, conceptual synthesis, and policy proposals. This distinction is essential because the manuscript's strongest contribution is the proposed integration, not proof that the sequence already operates as described.

- Addressed through the softened causal language and the explicit observational-data caveat on confirmation bias described above, both of which now separate established evidence from the conceptual synthesis.

Figures, Formatting, Citations, and Editorial Consistency:

- Figure 1 in Halo Effect section: The figure is relevant, but the image quality is low. Replace it with a clearer version if permission allows, or recreate the figure based on the reported data with proper attribution.
 - The original Figure 1 is retained with its citation; permission and resolution will be confirmed with the source ahead of production. A second, original schematic ("Proposed framework") was added in the Discussion to reduce reliance on reproduced figures.
- Figure 1 caption: Use a consistent citation format and change "Pg" to the journal's preferred page notation.
 - "Pg" replaced with "p." throughout figure captions and quotations.
- Section numbering: The manuscript has duplicate and incorrect Roman numerals. Renumber sections after all structural revisions are complete and ensure the roadmap matches the final numbering.
 - Renumbered I–VIII throughout, with the Introduction's roadmap paragraph corrected to match.
- Citation style: The manuscript mixes author-date citations with numbered bibliography markers and stray superscript numbers. Choose one style and apply it consistently.
 - Standardised to a single numbered bibliography (1)–(48) with matching in-text author-date citations; stray superscript numerals removed from the body text.
- Spelling convention: British and American spellings are mixed, including behavioural/behavioral, organisation/organization, judgement/judgment, neighbourhood/neighborhood, and maximise/maximize. Select one convention based on journal requirements.
 - Standardised to British spelling throughout (behavioural, organisation, judgement, neighbourhood, maximise, analysing).
- Punctuation: Several dash types appear throughout the paper, and some have encoding problems. Replace them with standard punctuation and review spacing around all dashes, parentheses, and citations.
 - Standardised punctuation and spacing around dashes, parentheses, and citations.
- Numbers and units: Standardize percentages, currency, PM_{2.5} notation, µg/m³, euro symbols, and CO₂ notation. Ensure spaces and superscripts follow one style.

- PM2.5 concentrations standardised to $\mu\text{g}/\text{m}^3$ with explicit WHO threshold comparisons; currency and percentage formatting standardised.
- Quotations: Verify all direct quotations against the original sources and include page numbers consistently.
 - Page numbers added/standardised across all direct quotations (e.g., Nisbett and Wilson 1977, p. 1; Nickerson 1998, p. 1; Kollmuss and Agyeman 2002, p. 249; King 2014, p. 21; Kahneman and Tversky 1981, p. 454), and the Kollmuss/Kahneman-Tversky wording checked against source.
- Source accuracy: Several years and author names differ between the manuscript and bibliography. Cross-check every in-text citation against the reference list.
 - Corrected discrepancies: Lord et al. (1979, not 1981); Schwartz (2004, not 2024); Karle et al. (2015, not 2013).
- Bibliography: Although the references were not reviewed line by line here, the list should be standardized for capitalization, author formatting, dates, URLs, DOI presentation, and hanging indentation before publication.
 - Bibliography rebuilt as a consistently formatted, numbered reference list (1)–(48), with several new supporting references added (He et al., 2018; Liu et al., 2018; Boso et al., 2025; Casillia and Hickey, 2016; Weiler et al., 2022; Hurst and Stern, 2020; Ditto et al., 2017; Washburn and Skitka, 2017; Ballew et al., 2019; Shen et al., 2023; Michelsen and Madlener, 2013; Faries, 2016).

I again thank both reviewers for their careful reading and appreciate the thoughtful feedback that has helped to strengthen the manuscript.

Final Review: Submission 100172

Thank you to the author for the considerable time and effort devoted to revising this manuscript. It is clear that the previous feedback was taken seriously, and I appreciate the thoughtful way in which many of the concerns raised during the initial review have been addressed. Overall, I believe the manuscript has been meaningfully strengthened and is much closer to publication readiness.

One of the most important improvements is the treatment of the proposed relationship among the halo effect, confirmation bias, and loss aversion. In my initial review, I felt that this was potentially the manuscript's most original contribution, but that the distinction between an interesting conceptual framework and an empirically demonstrated interaction needed to be clearer. The revised Discussion now handles this distinction much more carefully. I particularly appreciate the acknowledgement that alternative sequencing is possible, the clearer recognition that the cited cases do not prove the proposed interaction, and the addition of suggestions for how the framework might be tested empirically in future research.

The Methodology section has also improved. The author now more clearly identifies the work as a narrative and interpretive review, provides the databases and search period, describes the search terms, and offers additional explanation regarding study selection and thematic organization. These additions make the review process more transparent and help the reader better understand the scope of the evidence being synthesized.

Before publication, however, I recommend one final careful editorial and technical review. There are still places where wording from the revision process appears to remain alongside replacement wording, creating duplicated or grammatically incomplete sentences. Typographical errors, section references, figure numbering, citation details, and bibliography consistency should also be checked carefully throughout the final clean manuscript.

I would additionally encourage the author to review the loss-aversion section once more for technical precision. The distinction between loss aversion and risk aversion is important, and some of the descriptions of the experimental choices, percentages, and interpretations remain difficult to follow. These should be checked directly against the original cited studies so that the terminology and numerical descriptions accurately represent the source findings.

Finally, the revised manuscript does a better job of acknowledging alternative explanations and the ethical and equity implications of the proposed policy applications. I encourage the author to preserve this caution in the final version, particularly where observational evidence is being used to discuss possible cognitive mechanisms or where an original policy application has not itself been empirically tested.

Overall, I am pleased with the progress made since the initial submission. The author has responded thoughtfully to the substantive concerns raised in the first review, and the manuscript's interdisciplinary approach and conceptual synthesis remain important strengths. The remaining issues appear manageable through a focused final revision and, in my view, do not require another full round of peer review.

Final Recommendation: Accept, subject to minor revisions.

The author has done a thorough revision and done a good job responding to the concerns of the previous round of reviews. The paper better distinguishes between evidence of heuristics and the proposed integrative framework. The discussion of the interaction between the halo effect, confirmation bias, and loss aversion has been aptly softened such that it is now presented as conceptual and acknowledges the need for future work to test the proposed model. These changes alone have considerably strengthened the paper. I also appreciate the additional discussion of ethical considerations surrounding the proposed policy interventions. The paper acknowledges interventions based on social status or identity-based communication may lead to negative, unintended consequences for low SES or disadvantaged peoples.

As the major concerns of my previous review have been addressed, **I recommend acceptance with minor revisions (mentioned below)** - the remaining revisions are mostly editorial/clarifications rather than needing another round of review.

First, the manuscript could use another round of careful proofreading. There are grammatical and typographical errors that remain in the revised version. Some examples:

- multiple
- to in air quality perception
- hypothetical example
- Knobloch **at** al.
- starkly documentations
- As cited; wood-burning stoves
- that is larger than the sum of the impac of each heuristics sre complements

Additionally, the proposed experimental test of the interaction should be described more precisely. The manuscript mentions a “full factorial 3x3 experimental design” but if the intention is to manipulate the presence/absence of three heuristics, I believe a 2x2x2 factorial design would be appropriate. Perhaps I am misunderstanding, but it could use clarification either way. Or the paper should explain what three levels are intended for each heuristic. Also, later in the section (last paragraph) you propose a factorial test that involves only two heuristics. It would be helpful to clarify why the proposed test is limited to two heuristics if the broader framework involves all three. Additionally, as you mention, a factorial design is appropriate for testing if multiple manipulated factors interact. However, since you propose the relationship might be sequential, it may be useful to suggest a design that could distinguish an interaction from a sequential pathway. The authors could consider proposing an experiment that measures intermediate processes (eg. confirmation bias related belief updating) after manipulation of the halo effect, and then assesses if this alters the effect of loss aversion on behavior.

Finally, at the end of Section 1, you give the roadmap of the sections. You end with “Section VII concludes”. Section VIII is the conclusion, whereas Section VII is your discussion where you introduce your proposal – which is a considerable contribution of your paper. I would be sure to mention explicitly in your section description that you will integrate all three heuristics in a framework.