Andrea Verzobio[a], Denise Bolognani[b], John Quigley[c] and Daniele Zonta[a,b]*

[a]Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, Scotland, UK; [b]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy; [c]Department of Management Science, University of Strathclyde, Glasgow, Scotland, UK

**corresponding author: Daniele Zonta, Department of Civil, Environmental and Mechanical Engineering, University of Trento, Via Mesiano 77, 38123, Trento, Italy.*
E-mail address: daniele.zonta@unitn.it; daniele.zonta@strath.ac.uk.

Email address of all co-authors:
andrea.verzobio@strath.ac.uk; denise.bolognani@unitn.it; j.quigley@strath.ac.uk; daniele.zonta@unitn.it

Consequences of representativeness bias on SHM-based decision-making

Judging the state of a bridge based on SHM observations is an inference process, which should be rationally carried out using a logical approach. However, we often observe that real-life decision makers depart from this ideal model of rationality, judge and decide using common sense, and privilege fast and frugal heuristics to rational analytic thinking. For instance, confusion between condition state and safety of a bridge is one of the most frequently observed examples in bridge management. The aim of this paper is to describe mathematically this observed biased judgement, a condition that is broadly described by Kahneman and Tversky's representativeness heuristic. Particularly, we examine how this heuristic affects the interpretation of data, providing a deeper understanding of the differences between a method affected by cognitive biases and the classical rational approach. Based on the literature review, we identify three different models reproducing an individual behaviour distorted by representativeness. We apply these models to the case of a transportation manager who wrongly judges a particular bridge unsafe simply because deteriorated, regardless its actual residual load-carrying capacity. We demonstrate that application of any of the three heuristic judgment models correctly predicts that the manager will mistakenly judge the bridge as unsafe based on the observed condition state. While we are not suggesting in any way that representativeness should be used instead of rational logic, understanding how real-life managers actually behave is of paramount importance when setting a general policy for bridge maintenance.

Keywords: heuristics; representativeness; Bayesian inference; decision-making; reliability; bridge management

1. Introduction

Structural health monitoring (SHM) is commonly recognized as a powerful tool that allows bridge managers to make decisions on maintenance, reconstruction and repair of their assets (Bolognani, et al., 2018). The logic of making decisions based on SHM is formally stated in Cappello, Zonta and Glisic (2016), under the assumption that the decision maker is an ideal rational agent, who judges using Bayes' theorem (Bolstad,

2010) and decides consistently with Neumann-Morgenstern's Expected Utility Theory (EUT) (Neumann & Morgenstern, 1944). Not that surprisingly, it is often observed that real-life decision makers depart from this ideal model of rationality, judging and deciding using common sense and preferring fast and frugal heuristics to rational analytic thinking. Hence, if we wish to describe mathematically and to predict the choices of real-world bridge managers, we have to accept that their behaviour may not necessarily be fully rational. Biased judgement and decision-making have been widely reported and systematically investigated since the 1970s in the fields of cognitive sciences, social sciences and behavioural economics: key papers include the fundamental works by Kahneman and Tversky (Kahneman & Tversky, 1973; Kahneman & Tversky, 1979; Tversky & Kahneman, 1974; Tversky & Kahneman, 1983); Kahneman's famous textbook (Gilovich, et al., 2002) is an extensive reference for those approaching the topic for the first time.

As regards SHM-based bridge management, apparent irrational behaviours are reported in Zonta, Glisic and Adriaenssens (2014), Bolognani et al. (2017), Tonelli et al. (2018), Verzobio et al. (2019), and also suggested in Cappello et al. (2016). In particular, a typical example of cognitive bias frequently observed in bridge management is the confusion between condition state and safety of a bridge, as reported for instance in Zonta, Zandonini and Bortot (2007). Safety is about the capacity of a bridge to withstand the traffic loads and the other external actions without collapsing, while the condition state expresses the degree of deterioration of a bridge, or bridge element, with respect to its design state. The condition state is usually apprised through a combination of routing visual inspections, non-destructive evaluation and SHM. It is expressed in the form of a condition index that depends on the particular management system. For example, bridge management systems based on AASHTO (1997) Commonly Recognized (CoRe)

Standard Element System, such as PONTIS, BRIDGIT and the APT-BMS reported in Zonta et al. (2007), classify the state of an element on a scale from 1 to 5, where 1 means 'as per design' and 5 corresponds to the most severe observable deterioration state. On the contrary, the safety of a bridge is typically encoded in its probability of failure P_F, reliability index β, or safety factor γ, evaluated through formal structural analysis.

Condition state and safety are obviously correlated (logically, the load-carrying capacity of a deteriorated bridge is equal or lower than that of the same bridge in undamaged condition) but are not the same. For example, an old bridge can be unsafe, regardless of its preservation state, simply because it was designed to an old code, which does not comply with the current load demand. As a counterexample, we may have the case of a bridge, severely deteriorated, but still with enough capacity to safely withstand all the external loads, either because of overdesign or simply because its deterioration does not affect its load-carrying capacity. In principle, rational bridge management should target the safety of the bridge stock, and therefore prioritize retrofit of unsafe bridges, regardless of their degree of deterioration. In practice, it is frequently observed that bridge managers tend to delay retrofit of substandard bridges which do not show sign of deterioration, while repair promptly deteriorated bridges as soon as the damage is observed, regardless of the actual residual load-carrying capacity. The biased rationale behind this apparent behaviour is that undamaged bridges 'look' safe, while damaged bridges 'look' unsafe, simply because, generally speaking, it is acknowledged that deterioration negatively affects safety.

The aim of this paper is to describe mathematically this observed biased judgement, a condition that we will show, is broadly described by Kahneman and Tversky's representativeness heuristic (Kahneman & Tversky, 1972). We clarify that it is not an objective of this paper to suggest that it is correct to use representativeness to judge the state of a bridge: we presume it is evident to any reader that we shall always judge using rational logic, rather than a heuristic method. Indeed, our goal here is to verify whether the irrational judgment sometimes observed in bridge managers could be described and possibly predicted using Kahneman and Tversky's representativeness heuristic model. Being able to predict the behavior of an irrational manager is necessary when we set a general policy for bridge maintenance and we know, or suspect, that someone else who is going to enact the policy may behave irrationally. As an example, Gong and Frangopol (2020) discuss a case where modelling the irrational behavior of a manager is instrumental to an optimization process in bridge maintenance, and use Kahneman and Tversky's Prospect theory (Kahneman & Tversky, 1979) to simulate the biased decision of bridge managers. The authors conclude that the optimal maintenance policy should change if we properly account for the heuristic behavior of the decision makers.

To address our goal, we begin, in Section 2, with a review of the formal framework of rational decision based on SHM information. Section 3 discusses various classical judgmental heuristics and the consequential biases, while, in Section 4, various mathematical models of representativeness are analysed and formulated to appropriately reproduce the heuristic behaviour. In Section 5, a classical representativeness problem is discussed to assess these models. Finally, Section 6 presents an engineering application where the model is used to reproduce the biased evaluation about the safety of a bridge, based on the condition state apprised through visual inspections. Some concluding remarks are presented at the end of the paper.

2. SHM-based decision-making rational framework

We refer to the problem of optimal decision-making based on data provided by SHM. As shown in Figure 1, SHM-based decision-making is properly a two-step process, which includes the judgement of the state of the structure h based on the observations $\mathbf{y}$, and the decision of the optimal action a_{opt} based on the uncertain knowledge of the state. Within the scope of this paper, we define observation to be any information acquired on site which is suitable to infer the state of the structure. Sources of observation, in the broad sense, could be visual inspections, site tests, sensors temporarily or permanently installed on the structure.

[Figure 1. The rational process of SHM-based decision-making]

Assume that the safety state of the bridge is described by one of n mutually exclusive and exhaustive state hypothesis $\mathcal{H} = \{h_1, h_2, \ldots, h_j, \ldots, h_n\}$ (e.g.: h_1 = 'safe', …, h_n = 'failure'). Further assume that observing the bridge, or bridge element, either through visual inspection or SHM, ultimately consists of assessing its condition out of a number of m possible classes $\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_i, \ldots, \mathrm{C}_m$ which express its degree of damage or deterioration (e.g.: C_1 = 'not damaged', C_2 = 'moderately damaged', C_3 = 'severely damaged', …). Therefore, the value of an observation y_i is one of the possible condition classes: $y_i \in \{\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \mathrm{C}_4, \mathrm{C}_5\}$. Multiple independent observations on the same bridge may occur because of repeated inspections by different inspectors, or redundant independent measurements by the monitoring system. We indicate with vector $\mathbf{y}$ the full set of observations $\mathbf{y} = \{y_1, y_2, \ldots, y_k \ldots, y_N\}$. The likelihood of condition C_i for a bridge, or bridge element, in state h_j is then encoded in the probabilistic distribution $\mathrm{P}(\mathrm{C}_i | h_j)$.

If we restrict the problem to a single-observation case, the first step of the process consists of judging the state of a structure h_j based on the i-th class observed C_i. In the presence of uncertainty, the state of the structure after observing the class C_i is probabilistically described by the posterior probability $\mathrm{P}(h_j | \mathrm{C}_i)$, and the inference process followed by a rational agent is mathematically developed in Bayes' rule (Bolstad, 2010; Sivia & Skilling, 2006):

$$\mathrm{P}(h_j|\mathrm{C}_i) = \frac{\mathrm{P}\left(\mathrm{C}_i \middle| h_j\right)\mathrm{P}(h_j)}{\mathrm{P}(\mathrm{C}_i)}, \tag{1}$$

where $\mathrm{P}(h_j|\mathrm{C}_i)$ is the posterior knowledge of the structural state and represents the best estimation after the acquisition of SHM observation. It depends on the likelihood $\mathrm{P}\left(\mathrm{C}_i \middle| h_j\right)$ and the prior knowledge $\mathrm{P}(h_j)$, which is our estimate of the structural state h_j before the acquisition of the observation. $\mathrm{P}(\mathrm{C}_i)$ is simply a normalization constant, referred to as evidence, calculated as:

$$\mathrm{P}(\mathrm{C}_i) = \sum_{j=1}^{n} \mathrm{P}\left(\mathrm{C}_i \middle| h_j\right)\mathrm{P}(h_j). \tag{2}$$

The second step of the process starts after the assessment of the posterior probability of the structure, and concerns choosing the 'best' action. The decision maker can choose between a set of M alternative actions a_1, a_2, …, a_M (e.g.: a_1 = 'do nothing', a_2 = 'limit traffic', a_3 = 'close the bridge to traffic', …). Taking an action produces measurable consequences (e.g.: a monetary gain or loss, a temporary downtime of the structure, in some case causalities) and the consequences of an action can be mathematically described by several parameters (e.g.: the amount of money lost, the number of days of downtime, the number of casualties), encoded in an outcome vector $\mathbf{z}$. The outcome $\mathbf{z}$ of an action depends on the state of the structure; thus, it is a function of both action a and state h_j, i.e. $\mathbf{z}(a, h_j)$. When the state is certain the consequence of an action is deterministically known; therefore, the only uncertainty in the decision process is the state of the structure h_j. The rational decision-maker ranks actions based on the consequences $\mathbf{z}$ through a utility function $U(\mathbf{z})$, which can vary among different individuals with different behaviors. According to the different risk appetite of the decision-maker, the utility function can be risk neutral, risk adverse or risk seeking. Expected utility theory (EUT) describes the analysis of decision-making under risk and is considered as a normative model of rational choice (Parmigiani & Inoue, 2009). EUT was introduced by von Neumann and Morgenstern (1944) and later developed in the form that we currently know by Raiffa and Schlaifer (1961). Its axioms state that the decision-maker ranks their preferences based on the expected utility *u*, defined as:

$$u(a) = E_{h_j}\left[U\left(\mathbf{z}(a, h_j)\right)\right], \tag{3}$$

where E_{h_j} is the expected value operator of the random variable h_j, while U indicates the utility function. The latter is very important and represents the evaluation of a decision-maker's beliefs about the outcome $\mathbf{z}$. The decision-maker then chooses the action that maximizes the expected utility.

In summary, the rational way to decide based on observation in the presence of uncertainties goes through a judgment based on Bayes' theorem and a proper decision based on EUT.

3. Heuristics and biases

While an ideal rational decision maker judges and decides using Bayes' theorem and EUT, it is frequently observed that most people in everyday life favor heuristic approaches (Gilovich, Griffin & Kahneman, 2002; Kahneman & Tversky, 1979) to this rational framework in order to judge or make decisions.

The concept of heuristic has been defined in different ways in the scientific literature, depending on the discipline and the scope of application, see for instance Tonge (1960), Feigenbaum and Feldman (1963), Romanycia and Pelletier (1985), Gigerenzer and Gaissmaier (2011). For the purpose of this paper, we define a heuristic, together with Feigenbaum and Feldman (1963), as any approach to judgement or decision based on rules of thumb, logical simplifications or shortcuts rather that the proper rational process, as described in Section 2. Possibly, the most important contribution to the formal characterization of the heuristic behavior is the work that Kahneman and Tversky carried out in the early 1970s (Kahneman & Tversky, 1972; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974), which had a significant impact to the understanding and description of the human behavior and represents the basis of a discipline we currently refer to as *behavioral economics*. They developed the so-called *heuristics and biases approach*, challenging the dominance of strictly rational models. The main innovation lays in the analysis of the descriptive adequacy of ideal models of judgment and in the proposal of a cognitive alternative that explained human error without invoking motivated irrationality. Evidence displays that people's assessments of likelihood and risk do not conform to the laws of probability. They offer a list of frequently observed heuristics (Tversky & Kahneman, 1974), which include:

(1) Representativeness. Events are ranked according to their representativeness; people consistently judge the more representative event to be the more likely, whether it is or not (Kahneman & Tversky, 1972). Representativeness is not affected by several factors that affect rational judgments instead and this leads to relevant biases, such as: insensitivity to prior probability, insensitivity to sample size, misconceptions of chance, insensitivity to predictability, illusion of validity and misconceptions of regression (Tversky & Kahneman, 1974).

(2) Availability. An individual evaluates the frequency of classes or the probability of events by availability, i.e. by the ease with which relevant instances come to mind (Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). Thus, a person

could estimate the numerosity of a class, the likelihood of an event or the frequency of co-occurrences by assessing the ease with which the relevant mental operation of retrieval, construction or association can be conducted. It leads to predictable biases, e.g.: biases due to the retrievability of instances, biases due to the effectiveness of a search set, biases of imaginability and biases in the judgment of the frequency with which two events co-occur, i.e. illusory correlation.

(3) Adjustment or anchoring. People make estimates by starting from an initial value (which may be suggested by the formulation of the problem, or it may be the results of a partial computation), that is adjusted to yield the definitive answer. However, adjustments are typically insufficient, that is, different starting points yield different estimates, which are biased toward the initial values, and this phenomenon is called anchoring (Tversky & Kahneman, 1974).

Depending on their nature, a heuristic can affect the process outlined in Section 2 in the inference step, in the decision step, or in both cases. The rest of the paper will focus on the representativeness, as the heuristic that better reproduces the irrational behaviour introduced in Section 1. This specific heuristic affects the inference step of the process, i.e. the judgment.

4. The representativeness heuristic

Representativeness is commonly intended as the level of how well or how accurately something reflects upon a sample. A judgment is biased by the representativeness heuristic when the ordering of hypotheses h_j by subjective perceived probabilities coincides with their ordering by representativeness, rather than by Bayes' posterior probability (Kahneman & Tversky, 1972). In other words, a hypothesis, or event A, is judged more probable than a hypothesis, or event B, whenever A appears more representative than B. Citing Kahneman and Tversky (1972), an individual who follows the representativeness heuristic "*evaluates the probability of an uncertain event, or a sample, by the degree to which it is: (i) similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated*". This criterion for assessment does not coincide with the Bayesian posterior assessment and so results in a bias.

The literature illustrates numerous cases of behavioral experiments where representativeness bias is observed. For example, in a classic experiment reported in Tversky and Kahneman (1974), the interviewee is asked to assess the probability of Steve's employment from a list of possibilities (e.g. farmer, salesman, airline pilot, librarian or physician), simply based on this description: "*Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail*." It is observed that most interviewees tend to judge highly likely that Steve is a librarian, simply because the description provided is representative of the stereotype of a librarian, and with complete disregard for the proportion of the population that are librarians compared with the other employments. This example also clarifies that to be representative an uncertain event should not only be similar to its parent population, but it should also reflect the properties of the uncertain process by which it is generated. This agreement on the representativeness formulation is in line with the definition in Tversky and Kahneman (1983); they write that: "*an attribute is representative of a class if it is very diagnostic; that is, the relative frequency of this attribute is much higher in that class than in the relevant reference class.*"

While representativeness heuristic has been widely analysed from a descriptive point of view, in the literature there are only few models attempting to describe this heuristic from a mathematical perspective, see for instance Edward (1968), Grether (1980; 1992), Gigerenzer (1995), Barberis, Shleifer and Vishny (1998), Tenenbaum and Griffiths (2001), Bordalo, Coffman, Gennaioli and Shleifer (2016). While introducing these models, we have two key research questions to explore the definition of representativeness and its application, which are:

(1) What is the mathematical formulation of representativeness proposed by different authors?
(2) To what extent and how does representativeness bias the final judgment in comparison to Bayes' rule?

4.1 Formulation of Representativeness

In the literature mentioned above there is a general agreement whereby the degree of representativeness of an observable class C_i for a reference hypothesis h_j is in some way related to the odds of observable C_i, which is the ratio between its likelihood $\mathrm{P}(\mathrm{C}_i|h_j)$ and the likelihood of its negation $\mathrm{P}(\mathrm{C}_i|-h_j)$, where $-h_j$ denotes the set of alternative hypotheses.

Edward (1968), Gigerenzer (1995) and Bordalo et al. (2016) all define the quantity representativeness $\mathrm{R}(\mathrm{C}_i, h_j)$ of a class C_i for the reference hypothesis h_j, exactly as the odds of class C_i:

$$\mathrm{R}(\mathrm{C}_i, h_j) = \frac{\mathrm{P}(\mathrm{C}_i|h_j)}{\mathrm{P}(\mathrm{C}_i|-h_j)}. \tag{4}$$

Therefore, they assume that a class C_i is representative for a hypothesis h_j, relative to an alternative hypothesis $-h_j$, if it scores high on the likelihood ratio described by Equation (4).

Similarly, Tenenbaum and Griffiths (2001) define representativeness with the likelihood ratio described by Equation (4), but using a logarithm scale, apparently to provide a more natural measure of how good a class C_i is in representing a hypothesis h_j:

$$\mathrm{R}(\mathrm{C}_i, h_j) = \log \frac{\mathrm{P}(\mathrm{C}_i | h_j)}{\mathrm{P}(\mathrm{C}_i | -h_j)}. \tag{5}$$

Grether (1980; 1992) agrees on Equation (5) for a problem with two possible hypotheses. In the case of more alternative hypotheses, Tenenbaum and Griffiths (2001) suggest the following expression:

$$\mathrm{R}(\mathrm{C}_i, h_j) = \log \frac{\mathrm{P}(\mathrm{C}_i | h_j)}{\sum_{h_k,\, \mathrm{k} \neq \mathrm{j}} \mathrm{P}(\mathrm{C}_i | h_k)\, \mathrm{P}(h_k | -h_j)}, \tag{6}$$

where $\mathrm{P}(h_k | -h_j)$ is the prior probability of the k-th hypothesis, given that the reference hypothesis h_j is not the true explanation of C_i: 0 when $j = k$ and $\mathrm{P}(h_k)/(1-\mathrm{P}(h_j))$ when $j \neq k$. Equation (6) effectively says that C_i is representative of h_j to the extent that its likelihood under h_j exceeds its average likelihood under alternative hypotheses.

4.2 Representativeness in judgment

Before revising the mathematical models proposed to reproduce the representativeness bias in judgment, recall that the rational way to judge the probability of a hypothesis h_j based on an observation class C_i is to calculate its posterior probability $\mathrm{P}(h_j | \mathrm{C}_i)$ in Bayesian sense, using Equation (1). When judging using the representativeness heuristic, an individual ranks the hypothesis h_j by a subjective perceived probability which departs from standard Bayesian posterior. In analogy with Bordalo et al. (2016), we define this subjective perceived probability as distorted posterior $\mathrm{P}(h_j|\mathrm{C}_i)^{st}$. While all authors agree on that representativeness distorts judgment, there is not a general agreement on the cognitive mechanism whereby representativeness affects the distorted posterior probability, i.e. how the standard Bayes' rule, which reflects the judgment of a rational thinker, must be adjusted to consider representativeness instead. Most authors do not provide an explicit expression for the distorted posterior, but understand the vanilla statement that ordering hypotheses by perceived probability follows representativeness rather than Bayesian posterior. From a strict mathematical standpoint, it is possible to define different models of distorted posterior that satisfy this statement. A simple approach would be through assuming that (i) representativeness is used instead of likelihood and (ii) the prior information is neglected. In this case, judgment by representativeness should be consistent with the following expression:

$$\mathrm{P}\left(h_j \middle| \mathrm{C}_i\right)^{st} = \frac{\mathrm{R}(\mathrm{C}_i, h_j)}{\mathrm{R}(\mathrm{C}_i, h_j) + \mathrm{R}(\mathrm{C}_i, -h_j)}. \tag{7}$$

Some of the authors introduced above provide more refined models. Bordalo et al. (2016) suggest that representativeness $\mathrm{R}\left(\mathrm{C}_i, h_j\right)$ distorts Bayesian likelihood $\mathrm{P}\left(\mathrm{C}_i, h_j\right)$ as follows:

$$\mathrm{P}\left(\mathrm{C}_i \middle| h_j\right)^{st} = \mathrm{P}\left(\mathrm{C}_i \middle| h_j\right) \cdot \left(\mathrm{R}(\mathrm{C}_i, h_j)\right)^{\theta}, \tag{8}$$

where $\theta \geq 0$ is a subjective parameter that describes how heavily representativeness biases the likelihood. According to the same authors, this parameter should be calibrated with cognitive tests and could vary considerably among different people. A biased posterior is therefore inferred, using this distorted likelihood into Bayes' theorem:

$$\mathrm{P}\left(h_j \middle| \mathrm{C}_i\right)^{st} = \frac{\mathrm{P}\left(\mathrm{C}_i \middle| h_j\right)^{st} \mathrm{P}(h_j)}{\mathrm{P}(\mathrm{C}_i)^{st}}, \tag{9}$$

where $P(C_i)^{st}$ is the distorted evidence, calculated as:

$$P(C_i)^{st} = \sum_{j=1}^{n} P\left(C_i \middle| h_j\right)^{st} P(h_j). \tag{10}$$

It is easily noticed that Equation (9) is exactly Bayes' theorem when $\theta = 0$.

A different approach is provided by Grether (1980; 1992). The author suggests a model that provides the final judgment of h_j, by considering the representativeness heuristic:

$$\log O\left(h_j \middle| C_i\right) = \alpha + \beta_1 \cdot R\left(C_i, h_j\right) + \beta_2 \cdot \log O(h_j), \tag{11}$$

where $O(h_j|C_i)$ is the posterior odds, $R\left(C_i, h_j\right)$ is the representativeness calculated as in Equation (5), $O(h_j)$ is the prior odds, while α, β_1 and β_2 are subjective parameters that must be calibrated. Thus, the interpretation of Kahneman and Tversky's representativeness heuristic suggested by the author is that individuals place greater weight on the likelihood ratio than on the prior odds. Consequently, the author proposed $\beta_1 > \beta_2 \geq 0$ for this inference model, in contrast with $\alpha = 0$, $\beta_1 = \beta_2 > 0$ of Bayes' rule.

With the aim to compare these last two judgement models, we express Bordalo et al.'s model, stated in Equation (8), in its logarithmic posterior odds:

$$\log O\left(h_j \middle| C_i\right) = (2\theta + 1) \cdot R\left(C_i, h_j\right) + \log O(h_j), \tag{12}$$

where $R\left(C_i, h_j\right)$ is, in the same way as in Equation (11), the representativeness calculated as in Equation (5). It is possible to notice that this final equation agrees with the one proposed by Grether, i.e. Equation (11), if it is assumed that $\alpha = 0$, $\beta_1 = (2\theta + 1)$ and $\beta_2 = 1$. This means that the two models are based on the same mathematical formulation, they only differ in the representation of the subjective parameters.

In summary, while there is a general agreement on the definition and the mathematical formulation of the representativeness, different inference models are proposed or understood to describe the biased judgment. Moreover, some of these models account for a number of subjective parameters that have to be properly calibrated on the individual who judges.

5. A classical representativeness problem

Before developing the bridge engineering problem that is motivating this research, in this section we discuss how the models introduced in Section 4 apply to a classical representativeness problem, reported in different forms in Tversky and Kahneman (1974), Griffin and Tversky (1992), Tenenbaum and Griffiths (2001), Griffiths and Tenenbaum (2007).

Consider two coin-flip sequences, C_1 = HHHHH and C_2 = HTTHT, where H is for *Head* and T for *Tail*. To start, we would like to clarify the difference between representativeness and likelihood. The first question we ask ourselves is: which of the two sequences is more *representative* for a fair coin? We presume that most of the readers would answer sequence C_2. Actually, we expect that a fair coin would generally produce a random sequence of H and T, as in C_2, while a sequence of H only, as in C_1, looks intuitively peculiar from a genuinely fair coin. Intuitively, we conclude that the representativeness of sequence C_2 is greater than the representativeness of sequence C_1 under the assumption of fair coin h_{FC}. In formula:

$$\mathrm{R}(\mathrm{C}_2, h_{\mathrm{FC}}) > \mathrm{R}(\mathrm{C}_1, h_{\mathrm{FC}}). \tag{13}$$

However, which of the two sequences is more likely to occur for a fair coin? In this case we can simply calculate the likelihood of a sequence, i.e. the probability of obtaining that particular sequence C_i conditional to the assumption of fair coin $\mathrm{P}(\mathrm{C}_i|h_{\mathrm{FC}})$, by computing the possible combinations. If the coin is fair, for each toss there is equal probability of p = 1/2 of H or T. Therefore, the particular sequence C_2, which is the result of 5-coin tosses, has the following likelihood:

$$P(C_2{=}\text{HTTHT}|h_{FC}) = \left(\frac{1}{2}\right)^5 = 0.0313. \tag{14}$$

Notice now that even sequence C_1 is a possible output of 5-coin tosses, and therefore its likelihood is exactly the same as C_2:

$$P(C_1{=}\text{HHHHH}|h_{FC}) = \left(\frac{1}{2}\right)^5 = 0.0313. \tag{15}$$

Let’s now ask to the layman the following question: a coin has produced the sequence C_1; based on this sequence, do you believe this coin is more likely to be fair or have a prevalence of H? Most of the interviewees, and possibly even the reader, answer that the coin is most likely unfair, i.e. with a prevalence of H. Let’s tackle the problem in logical terms using Bayes’ theorem. As regards the coin with a prevalence of H, we refer to a coin that mostly comes up heads with the term h_{MH}. In essence, the following posterior probability has to be calculated:

$$P(h_{FC}|C_1) = \frac{P(C_1|h_{FC})\ P(h_{FC})}{P(C_1|h_{FC})\ P(h_{FC}) + P(C_1|h_{MH})\ P(h_{MH})}. \tag{16}$$

The condition whereby it is more probable that the coin is fair is that the posterior probability of h_{FC} is greater than 0.5, or:

$$\frac{P(h_{FC})}{P(h_{MH})} > \frac{P(C_1|h_{MH})}{P(C_1|h_{FC})}, \tag{17}$$

which means that the ratio between the priors has to be greater than the ratio between the opposite likelihoods. If p is the probability of occurrence of H in a toss, *n* the number of coin tosses and *k* the number of H achieved in a sequence C_i, the likelihood of this sequence can be calculated as follows:

$$P(C_i | h_{FC}) = p^k (1 - p)^{n - k}. \tag{18}$$

In the case of a fair coin we have already observed that p = 1/2 and therefore $P(C_1 | h_{FC}) = P(C_2 | h_{FC}) = 0.0313$. On the other hand, coin h_{MH} is the one that mostly comes up heads, and therefore, since C_1 is a sequence with a prevalence of H, the only thing that it is possible to conclude is that:

$$P(C_1 | h_{MH}) > P(C_1 | h_{FC}). \tag{19}$$

Therefore, the only logical conclusion we can draw is that the ratio between the two likelihood, from Equation (17), is strictly greater than 1. In any case, it states nothing about the posterior because it depends also on the prior rate, i.e. how likely is a priori, before observing the sequence, that the coin is fair. There is always a value for the prior $P(h_{FC})$ whereby it is more probable that, given the sequence C_1, the coin is fair:

$$P(h_{FC} | C_1) > P(h_{MH} | C_1). \tag{20}$$

In conclusion, from a strict logical standpoint, the coin could be fair or nor fair depending on the prior information.

Let's make a numerical example: we assume we have a fair coin h_{FC}, and a coin that mostly comes up heads h_{MH}, assuming the probability of occurrence of H with this coin is p = 0.85. For concreteness, we choose the following prior probabilities for the two hypotheses: $P(h_{FC}) = 0.95$ and $P(h_{MH}) = 0.05$. First of all, all the likelihood and the representativeness values have to be calculated. Using respectively Equation (18) and Equation (4), we obtain:

$$P(C_1|h_{FC}) = 0.0313, \qquad P(C_2|h_{FC}) = 0.0313. \tag{21a,b}$$

$$P(C_1|h_{MH}) = 0.4437, \qquad P(C_2|h_{MH}) = 0.0024. \tag{22a,b}$$

$$R(C_1|h_{FC}) = 0.07, \qquad R(C_2|h_{FC}) = 13.04. \tag{23a,b}$$

$$R(C_1|h_{MH}) = 14.18, \qquad R(C_2|h_{MH}) = 0.08. \tag{24a,b}$$

These results confirm what we were presuming and clearly show the difference between representativeness and likelihood: while sequences HTTHT and HHHHH are equally likely for a fair coin, i.e. $P(C_1|h_{FC}) = P(C_2|h_{FC})$, the representativeness model shows that sequence HTTHT is clearly more representative for a fair coin than sequence HHHHH, i.e. $R(C_2, h_{FC}) > R(C_1, h_{FC})$. This outcome reflects the effect of such heuristic bias, because most people judge the sequence HTTHT to be more likely for a fair coin than sequence HHHHH, which does not appear random, even if the two sequences have the same probability of occurrence. The second column of Table 1 presents the achieved results.

Let's now calculate the Bayesian posterior probabilities that, given sequence C_1, the coin is fair, or it is the one that mostly comes up heads. Using Bayes' theorem, as in Equation (16), we achieve:

$$P(h_{FC}|C_1) = 57.27\% > P(h_{MH}|C_1) = 42.73\%. \tag{25}$$

Notice that, with the prior assumptions made, the rational conclusion is that the coin is most probably fair, even if it has yielded a sequence of 5 heads in a row. This result may sound counterintuitive to the layman, unfamiliar with formal logic, who tends to judge heuristically, driven by the representativeness of the observed result.

It is possible to reproduce this heuristic behaviour using, for instance, the vanilla inference model of Equation (7), which indeed yields the following distorted posterior judgments:

$$\mathrm{P}(h_{\mathrm{FC}}|\mathrm{C}_1)^{st} = 0.49\% < \mathrm{P}(h_{\mathrm{MH}}|\mathrm{C}_1)^{st} = 99.51\%, \tag{26}$$

which is to say that to the individual biased by representativeness, the coin looks most likely the one that mostly comes up heads. We find a similar result using the other inference model of Bordalo et al., as in Equation (9): using a subjective parameter θ = 0.8, which correspond to a high level of representativeness, we obtain:

$$\mathrm{P}(h_{\mathrm{FC}}|\mathrm{C}_1)^{st} = 1.88\% < \mathrm{P}(h_{\mathrm{MH}}|\mathrm{C}_1)^{st} = 98.12\%, \tag{27}$$

which again shows that a sequence of five heads heuristically (but mistakenly) suggests that the coin is the one that mostly comes up heads. Clearly, in this case the perceived posterior probability depends on the parameter θ, as will be discussed in detail in Section 6.3.

Table 1 presents the outcomes from all the inference models reviewed in Section 4, including Grether's, i.e. Equation (11), evaluated with $\alpha = 0$, $\beta_1 = 0.8$ and $\beta_2 = 0.2$, i.e. for an irrational manager with a high level of representativeness. It is evident that all the heuristic inference models, if evaluated using subjective parameters which correspond to a high level of representativeness, agree on judging most likely that the coin is the one that mostly comes up heads h_{MH}, in contrast to the rational conclusion inferred through Bayes' theorem.

Table 1. Achieved results for each model.

In conclusion, with this numerical example we have clarified the substantial difference between likelihood and representativeness. We have also shown how the representativeness bias may alter the posterior judgment to the point of suggesting conclusions opposite to those consistent with rational inference.

6. Case study

In this section we wish to verify whether the judgment models reviewed in Section 4 are suitable to describe the typical confusion between condition state and safety of a bridge frequently observed in bridge management. As described in Section 1, bridge managers often tend to delay retrofit of substandard bridges which do not show sign of deterioration, while repair promptly deteriorated bridges as soon as the damage is observed, regardless of their actual residual load-carrying capacity. We have already observed that the biased rationale behind this apparent behaviour is that undamaged bridges 'look' safe, while damaged bridges 'look' unsafe, simply because, generally speaking, we know that deterioration negatively affects safety.

We discuss this bias with reference to one of the case studies reported in Zonta et al. (2007), i.e. the SP65 bridge on the Maso River, which is operated by the Autonomous Province of Trento (APT). The bridge, shown in Figure 2(a), is a common type of bridge in the APT stock. The structure has two simple spans of 19.0 m and 22.0 m, and a total length of 43.0 m. Each span has four girders spaced at 2.1 m, 2.4 m and 2.1 m respectively. The cross-section of the girders is shown in Figure 2(b). The deck slab consists of 22–27 cm of reinforced concrete and a 15 cm surface layer of asphalt. The roadway width is 7 m with 0.70 m pedestrian pavements and hand railing on each side.

[Figure 2. SP65 bridge on Maso River: (a) overview; (b) plan view, elevation and cross section of the deck (Zonta, et al., 2007).]

Managing its bridges, APT uses an inventory model and condition state appraisal system consistent with the AASHTO (1997) Commonly Recognized (CoRe) Standard

Element System. The CoRe element standard has been adopted since 1995 by FHWA and AASHTO as broadly accepted way to represent bridges condition on a uniform scale. The CoRe element standard inventories a bridge into a set of Standard Elements (SE), each specified in term of quantity (surface, length or number). For example, the bridge deck of the SP65 bridge includes the following SE: slab, beam, pavement, sidewalk, guard rail and railing.

The state of deterioration of each element is appraised through routine visual inspections. The inspector classifies the state of deterioration of an element choosing among five possible deterioration levels, called Condition States (CS), specified, for each element type, in the inspector manual. Table 2 reports, as an example, the definition of the five CS of a concrete slab, or CoRe standard element #12, as reported in APT inspection manual available from the website of the APT (2018). As a general rule, Condition State 1 (CS_1) always means 'as per design', or 'no deterioration', while CS_5 corresponds to the most severe observable deterioration state.

While the deterioration condition is apprised through visual inspection, its safety level is evaluated separately, through a five-step formal assessment procedure (Zonta, et al., 2007), whose ultimate objective is to calculate the bridge reliability index β. We have already observed in the Introduction that condition state and safety are obviously correlated, but not the same thing, and that we can well have a severely deteriorated bridge which is perfectly safe or an intact bridge which is not safe. We have also noticed that a rational bridge manager should address safety above all, while in practice the intervention priority is often biased by the apparent state of deterioration of the bridges, regardless their actual residual load-carrying capacity.

Table 2. SE #12 concrete slab: state description for each Condition State (CS).

In this section, we want to numerically analyse and describe the following case:

- As far as its safety is considered, the bridge could be in two possible states: SAFE (h_S) or FAIL (h_F). SAFE means that, following to a formal safety assessment carried out by an expert structural engineer, the bridge load-carrying capacity is judged sufficient for the bridge to operate without restrictions. On the other hand, FAIL means that the bridge is not found to have sufficient load-carrying capacity and should be closed to traffic.
- Based on a frequentist analysis of the load-carrying capacity formally assessed for similar bridges of the same type and age, it is estimated that only one bridge out of one thousand is found to be in the FAIL state. We formalize this information assuming prior base rates $P(h_F) = 0.001$ for the state hypothesis FAIL, and therefore $P(h_S) = 0.999$ for the state hypothesis SAFE.
- Based on the last visual inspection, the bridge exhibits no or minimal deterioration, except for the concrete slab, which is classified in the most severe condition state, or CS_5.
- Based on the condition state assessed via visual inspection, the bridge manager judges the bridge in FAIL state.

This case study effectively describes a prototypical situation where the bridge manager judges the state of safety of the bridge based on the condition state of one of its elements and disregarding any information on its actual residual load-carrying capacity. The manager implicitly assumes that a severe deterioration of an element automatically implies that the bridge load-carrying capacity is insufficient, simply because deterioration is representative for a reduced capacity. We hypothesize this situation could be described as a case of the representativeness bias, where the safety is improperly judged based on how much deterioration is representative of loss in capacity.

In order to verify this conjecture, we will answer quantitatively the following questions:

(1) What is the likelihood $P(CS_5|h_F)$ of an unsafe bridge to be in CS_5?

(2) How much CS_5 is representative of a bridge in FAIL state?

(3) What is the proper posterior probability of this bridge to be in FAIL state?

(4) How does representativeness bias distort the manager judgment as to the bridge safety?

6.1 Likelihood and representativeness

To start, we have to define a proper likelihood distribution for each hypothesis, i.e. $P(CS_i|h_F)$ and $P(CS_i|h_S)$. In the following, the procedure used for the definition of the likelihood is the same as in Zonta et al. (2007).

According to (Melchers, 1999), we employ II level probabilistic methods, which allows to calculate the reliability index $\beta = -\Phi^{-1}(P_{h_F})$, where Φ is the cumulative normal distribution function. Two stochastic variables are considered: the loads effect S and the starting resistance R_0 of the bridge, both supposed to be Normal distributions (Norm), with their mean μ and coefficient of variation V. In formula:

$$f_{R_0}(r) = \text{Norm}(r, \mu_{R_0}, V_{R_0}), \quad f_S(s) = \text{Norm}(s, \mu_S, V_S). \tag{28a,b}$$

Because of the prioritization approach, we assume that the structure will not maintain its mechanical characteristics in the years, i.e. we have to take into account the deterioration of construction material through the following probabilistic degradation model (Zonta, et al., 2007):

$$R = R_0(1 - \delta(\mathrm{CS}_i)), \tag{29}$$

where $\delta(\mathrm{CS}_i)$ is a probabilistic capacity degradation function, depending only on the CS_i of the SE that controls the capacity of the structural unit at the limit state. Its density function δ_i is the probability density function of the loss in capacity when the element is in the i-th CS. Recall that the elements are rated based on visual inspections, δ_i represents the likelihood of a certain loss in capacity when the element has been rated into the i-th reference state. Typically, low values of CS, i.e. CS_1 CS_2 and CS_3, are not associated with any loss of capacity: in this case δ_i coincides with a Dirac delta function and therefore $R = R_0$. On the other hand, higher CS_i are associated with distributions that reflect the uncertainty of the system in correlating the actual loss in capacity, with the verbal description of the reference state proposed by the inspection manual. CS_4 is associated with a uniform distribution δ_4 of loss in capacity, for values of δ included in [0, 5%]. In the same way, the system associates the reference state 5 with a triangular distribution, for values of δ included in [5%, 70%], as Figure 3 shows.

[Figure 3. Capacity degradation function $\delta(\mathrm{CS}_i)$.]

Because most of the information required to define the distribution of capacity R and actions S are not explicitly contained in the system database, a simplified approach must be adopted. It is convenient to define a normalized capacity $r = R/\mu_S$, with mean value $\mu_r = \mu_{R_0}/\mu_S$, equal to the central safety factor γ_0, associated with the limit state Z, and a normalized demand $s = S/\mu_S$ with mean value $\mu_S = 1$. The coefficients of variations of the normalized variables γ and s are equal to those of R and S. The Normal distribution of the capacity and actions become:

$$f_{\gamma_0}(r) = \mathrm{Norm}(r, \gamma_0, V_R), \qquad f_S(s) = \mathrm{Norm}(s, 1, V_S), \tag{30a,b}$$

where the reliability index is related to the central safety factor γ_0 through the expression:

$$\beta = \frac{\gamma_0 - 1}{\sqrt{{V_R}^2 \cdot {\gamma_0}^2 + {V_S}^2}}. \tag{31}$$

Finally, the normalized limit state function is $z = r - s$, and the probability of failure P_{h_F} associated with the limit state Z coincides with that of z:

$$P_{h_F}(\mathrm{CS}_i) = P(Z < 0) = P(z < 0). \tag{32}$$

According to Eurocode 0, if we employ II level probabilistic methods, the target reliability index β for Class RC2 structural member in the Ultimate Limit State and with a reference time of 1 year is equal to $\beta = 4.75$. Assuming $V_R = 0.05$ and $V_S = 0.10$, from Equation (31) we can obtain $\gamma_0 = 1.96$. Once we know γ_0, the probability of failure $P_{h_\mathrm{F}}(\mathrm{CS}_i)$ is then calculated through Monte Carlo methods by computing the cumulative-time failure probability of the normalized limit state z, by using a normalized Gaussian distribution for the demand $f_S(r)$ and a normalized non-Gaussian distribution for the reduced capacity $r = \gamma_0(1 - f_\delta)$, which depends on CS_i:

$$f_r(r, \mathrm{CS}_i) = f_{\gamma_0}(r)(1 - f_\delta). \tag{33}$$

Consequently, assuming $P_{h_\mathrm{F}}(\mathrm{CS}_1) = P_{h_\mathrm{F}}(\mathrm{CS}_2) = P_{h_\mathrm{F}}(\mathrm{CS}_3)$, we obtain the following failure probability for each CS_i:

$$\left[P_{h_\mathrm{F}}(\mathrm{CS}_1);\ P_{h_\mathrm{F}}(\mathrm{CS}_2);\ P_{h_\mathrm{F}}(\mathrm{CS}_3);\ P_{h_\mathrm{F}}(\mathrm{CS}_4);\ P_{h_\mathrm{F}}(\mathrm{CS}_5)\right] = [2.35 \cdot 10^{-5};\ 2.35 \cdot 10^{-5};\ 2.35 \cdot 10^{-5};\ 6.61 \cdot 10^{-4};\ 2.04 \cdot 10^{-1}]. \tag{34}$$

Assuming the following a priori distributions for CS, i.e. $\mathrm{P(CS)} = [50\%, 20\%, 15\%, 10\%, 5\%]$, we can calculate the probability for each hypothesis, i.e. "FAIL = h_{F}" and "SAFE = h_{S}" respectively:

$$P_{h_{\mathrm{F}}} = \sum_{i=1}^{\mathrm{CS}_{max}=5} P_{h_{\mathrm{F}}}(\mathrm{CS}_i) \cdot P(\mathrm{CS}_i) = 0.0102, \tag{35}$$

$$P_{h_{\mathrm{S}}} = 1 - P_{h_{\mathrm{F}}} = 0.9898. \tag{36}$$

It is important to explain why these values are different from the assumed prior base rates (i.e. $\mathrm{P}(h_{\mathrm{F}})$ = 0.001 and $\mathrm{P}(h_{\mathrm{S}})$ = 0.999): APT, as most of the transportation agencies, calculates the nominal probability of collapse using a mechanical model, which is based on conservative assumptions and an estimate of the loss in capacity due to degradation, which is also based on conservative assumptions. In theory, if the model used by APT was unconditionally correct, then $P_{h\mathrm{F}}$ and the prior $\mathrm{P}(h_{\mathrm{F}})$ should be identical. However, in practice the model is clearly conservative and predicts a number of fail cases greater than 1% (i.e. $P_{h\mathrm{F}}$ = 0.0102), that is much higher than the actual number frequentistically observed (i.e. $\mathrm{P}(h_{\mathrm{F}})$ = 0.001). This is a very typical situation for transportation agencies because prediction models are deliberately conservative. In order to cope with this apparent contradiction, we assume that, although the model is overconservative in the prediction of the probability of collapse, the ratio between two different probability of collapse is reasonably correct. In other words, the mechanical model is not suitable to predict the actual absolute probability of collapse given the CS, but is reliable enough to predict that the probability of collapse of a structure in CS_5 is about 300 times bigger than the probability of collapse in CS_4.

Then, according to Bayes' rule, for both hypothesis "S = SAFE" and "F = FAIL" we can evaluate the relative likelihood distributions for each Condition State CS_i, as follows:

$$\mathrm{P}(\mathrm{CS}_i|h_\mathrm{F}) = \frac{P_{h_\mathrm{F}}(\mathrm{CS}_i)\cdot P(\mathrm{CS}_i)}{P_{h_\mathrm{F}}},\quad \mathrm{P}(\mathrm{CS}_i|h_\mathrm{S}) = \frac{(1-P_{h_\mathrm{F}}(\mathrm{CS}_i))\cdot P(\mathrm{CS}_i)}{P_{h_\mathrm{S}}}. \quad (37\text{a,b})$$

Figure 4 shows the results for each CS_i, which numerically correspond to the following likelihood distributions:

$$\mathrm{P}(\mathrm{CS}_i|h_\mathrm{S}) = [50\%, 20\%, 15\%, 10\%, 5\%],\quad \mathrm{P}(\mathrm{CS}_i|h_\mathrm{F}) = [0, 0, 0, 2\%, 98\%]. \quad (38\text{a,b})$$

[Figure 4. Likelihood distributions for each state hypothesis.]

After the evaluation of the likelihoods, we are interested in understanding how much CS_5 is representative of the bridge in FAIL state. We can calculate it according to Equation (4):

$$\mathrm{R}(\mathrm{CS}_5|h_\mathrm{F}) = 19.6, \qquad \mathrm{R}(\mathrm{CS}_5|h_\mathrm{S}) = 0.05. \quad (39\text{a,b})$$

These outcomes show that, as expected, CS_5 is very representative of the failure state of the bridge, with an enormous difference in comparison to the safe state of the bridge, i.e. $\mathrm{R}(\mathrm{CS}_5|h_\mathrm{F}) \gg \mathrm{R}(\mathrm{CS}_5|h_\mathrm{S})$: this is very important because we have learnt that this can be the reason for a distorted final judgment.

6.2 Posterior judgment

We evaluate the posterior judgment of the manager, in the case that the bridge is classified in CS_5. The proper posterior probabilities, computed using the rational framework provided by Bayes' theorem, results:

$$\mathrm{P}(h_\mathrm{F}|\mathrm{CS}_5) = 1.92\% < \mathrm{P}(h_\mathrm{S}|\mathrm{CS}_5) = 98.08\%. \quad (40)$$

This means that rational managers, in line with Bayes' rule and after observing CS_5, would judge the possibility that the bridge could be in the FAIL state as very unlikely.

However, we have introduced before that, based on the condition state assessed via visual inspection, the bridge manager has judged the bridge in FAIL state. It is possible to explain this judgment by evaluating the distorted posterior probability. Using the vanilla inference model of Equation (7), we achieve:

$$\mathrm{P}(h_\mathrm{F}|\mathrm{CS}_5)^{st} = 99.75\% > \mathrm{P}(h_\mathrm{S}|\mathrm{CS}_5)^{st} = 0.25\%. \quad (41)$$

Similarly, accepting the inference model of Bordalo et al., the distorted posterior probability is:

$$\mathrm{P}(h_\mathrm{F}|\mathrm{CS}_5)^{st} = 69.97\% > \mathrm{P}(h_\mathrm{S}|\mathrm{CS}_5)^{st} = 30.03\%. \quad (42)$$

In both cases the failure state turns out to be the most likely, and this outcome allows to explain the judgment of the manager, which is biased since CS_5 is very representative of a fault bridge.

Table 3 reports all the achieved results; the last row of the table presents again the results that come from the inference model of Grether, which agree with those obtained with the other biased models, i.e. the FAIL state is the most likely, in contrast to the rational conclusion inferred through Bayes' theorem.

Table 3. Achieved results for each model.

In summary, we have demonstrated that when an inspector judges the safety state of a bridge by only accounting for the observed condition state CS, they are biased by representativeness: in their posterior judgments they tend to neglect the prior probability of the failure condition, which is typically very low, $\mathrm{P}(h_\mathrm{F})$ = 0.001 in this specific case study, and to weight too much the ratio between the likelihood of the observations, which is the representativeness itself. Therefore, their final judgment results distorted in comparison to the one achieved by a manager who stick to rational thinking.

6.3 Discussion about inference models

To develop the numerical calculations in the previous sections, we had to assume specific values for the subjective parameters of the inference models introduced in Section 4.2, i.e. θ = 0.8, α = 0, β_1 = 0.8, β_2 = 0.2: these values correspond to a high level of the representativeness heuristic since they maximize the importance of R and minimise the contribute of the prior information. Since these parameters depend on different behaviour of people and could vary considerably, it is interesting to develop a sensitivity analysis in order to understand how they affect the model and then the conclusive results.

Let's take for instance the model of Bordalo et al.: as we can see from Equation (8), it depends only on one subjective parameter, i.e. $\theta \geq 0$. Figure 5 shows how the posterior failure probability of the bridge, after observing CS_5, varies according to θ: even if θ can be also larger than 1, we study just the interval $0 \leq \theta \leq 1$ since this is sufficient to understand how the results change. The previous assumption of θ = 0.8 resulted in $P(h_F|CS_5)$ = 69.97%, but we notice that the outcome is highly sensitive to the choice of θ: it changes from $P(h_F|CS_5)$ = 1.92% if θ = 0, i.e. in line with a rational manager who follows Bayes' rule, to $P(h_F|CS_5)$ = 88.49% if θ = 1, i.e. in line with an irrational manager biased with a high level of the representativeness heuristic. Furthermore, we observe that the posterior failure probability $P(h_F|CS_5)$ is larger than the posterior safe probability $P(h_S|CS_5)$ when $\theta > 0.67$. These results demonstrate the importance of calibrating properly the subjective parameters according to the specific inspector. The same generic conclusions can be extended to the model of Grether, since we have demonstrated that it is based on the same mathematical formulation.

[Figure 5. How the distorted posterior probability $P(h_F|CS_5)^{st}$ varies according to the subjective parameter θ.]

Conversely, the vanilla model introduced in Equation (7) is less sophisticated because it does not depend on a subjective parameter. Even if this may seem like a shortcoming, the results obtained in both Section 5 and Section 6 demonstrate the correctness of the vanilla model in reproducing the distorted judgment based on the representativeness bias. In detail, it is evident that its outcomes are similar to those that can be obtained assuming the maximum level of representativeness in the subjective parameters of the other inference models. As such, the vanilla model reproduces the behaviour of an inspector completely biased by this heuristic. This conclusion is consistent with the mathematical formulation of the model itself, since it overlooks the contribution of the prior and it completely replaces the likelihood with the representativeness.

7. Conclusions

Judging the state of a bridge based on SHM observations is an inference process which should be rationally carried out using a logical approach. However, we often observe that real-life decision makers depart from this ideal model of rationality, judge and decide using common sense, and privilege fast and frugal heuristics to rational analytic thinking. For instance, confusion between condition state and safety of a bridge is one of the most frequently observed examples in bridge management. In this contribution, we have demonstrated that this bias can be described by Kahneman and Tversky's representativeness heuristic.

A review of the technical literature shows that representativeness heuristic has been widely analysed from a descriptive point of view, while only few models have been proposed to describe this bias from a mathematical perspective. In the literature there is a general agreement that the degree of representativeness of an observable class for a reference hypothesis is in some way related to odds of observable quantities. However, there is not a general agreement on how the standard Bayes' rule, which is typically taken as the baseline model to reproduce the judgment of a rational thinker, should be distorted to consider representativeness. Most authors do not provide an explicit expression for the distorted posterior, but understand the statement that ordering hypotheses by perceived probabilities follows representativeness rather than Bayesian posterior. This is consistent with a distorted judgement model, here referred to as 'vanilla', whereby (i) representativeness is used instead of likelihood and (ii) the prior information is neglected. Bordalo et al. and Grether provide more refined models for reproducing the subjective distorted judgement, which allow to blend more flexibly likelihood, representativeness and prior information, through a number of subjective parameters, in order to better reproduce the distorted perception of a particular subject.

We have first applied these mathematical models to a classical literature representativeness problem, to better appreciate the difference among the various formulations of representativeness and heuristic judgement models. Next, we have applied the same models to the case of a transportation manager who wrongly judges a particular bridge unsafe simply because deteriorated, regardless its actual residual load-carrying capacity. Their judgment is biased due to the apparent behaviour that damaged bridges 'look' unsafe, in contrast with undamaged bridges which 'look' safe.

In the particular case study, we have demonstrated that Bayes' theorem correctly identifies the bridge as safe, while application of the three judgment models analysed (vanilla, Bordalo et al.'s and Grether's) all predict the manager will mistakenly judge the bridge as unsafe based on the observed condition state. Given the simplicity of the case study, which is essentially a two hypotheses inference problem where the individual distorted behaviour is characterized by the ordering of the two hypotheses by subjective probabilities, the three models are equivalent in this particular instance, as they reproduce equally well the observed distorted perception. The main difference between these three inference models is that 'vanilla' model reproduces the behaviour of an individual whose judgement is blatantly driven by representativeness, while the other two models describe more subtle forms of distorted judgment, whose limit cases are rational Bayesian inference on one side and the vanilla representativeness bias on the other. The three models may not be equivalent in a more complex setting, where the vanilla inference model may fail to reproduce the observed representativeness bias. Bordalo et al.'s and Grether's model are clearly more flexible, but at the same time very sensitive to a number of subjective parameters, which have to be accurately calibrated, typically with cognitive tests, on the particular individual whose distorted judgment is to be described. While it is not the objective of this paper, we suggest that there may be applications which require to identify precisely the best representativeness model, and then the subjective parameters, for example if we need to predict the rational behaviour of the manager in a future instance: in this case we would need additional observations of the manager behaviour, in order to identify the proper model. This can be done either by further unelicited observations or through proper cognitive tests in an elicitation process: see Verzobio et al. (2020) for an example of elicitation process applied to an engineering real-life case study.

To conclude, we reiterate once again that we are not suggesting in any way that representativeness should be used instead of rational logic. At the same time, predicting the actual behavior of managers is required when setting a general policy for bridge maintenance, acknowledging that the managers who are going to enact the policy may behave irrationally.

Declaration of interest statement

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

American Ass. State Highway and Transportation Off. (1997). *AASHTO guide for commonly recognized (CoRe) structural elements.* Washington D.C.: AASHTO.

Autonomous Province of Trento. (2018). *www.bms.provincia.tn.it/bms.* [Online].

Barberis, N., Shleifer, A. & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics,* 49(3), 307-343.

Bolognani, D., Verzobio, A., Tonelli, D., Cappello, C., Zonta, D. & Glisic, B. (2017). *An application of Prospect Theory to a SHM-based decision problem.* Proceedings of SPIE, Portland.

Bolognani, D, Verzobio, A., Tonelli, D., Cappello, C., Glisic, B., Zonta, D. & Quigley, J. (2018). Quantifying the benefit of strutural health monitoring: what if the manager is not the owner?. *Structural Health Monitoring,* 17(6), 1393-1409.

Bolstad, W. M. (2010). *Understanding Computational Bayesian Statistics.* NJ, USA: John Wiley & Sons.

Bordalo, P., Coffman, K., Gennaioli, N. & Shleifer, A. (2016). Stereotypes. *Quarterly Journal of Economics,* 131(4), 1753-1794.

Cappello, C., Zonta, D. & Glisic, B. (2016). Expected utility theory for monitoring-based decision-making. *Proceedings of the IEEE,* 104(8), 1647-1661.

Edward, W. (1968). Conservativism in Human Information Processing. In B. Kleinmuntz (Ed.), *Formal representation of Human Judgment* (17-52). New York: Wiley.

Feigenbaum, E. A. & Feldman, J. (1963). *Computers and thought.* New York: McGraw-Hill Inc.

Gigerenzer, G. (1995). How to improve Bayesian Reasoning Without Instructuion: Frequency Formats. *Psychological Review,* 102(4), 684-704.

Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62, 451-482.

Gilovich, T., Griffin, D. W. & Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment.* Cambridge University Press.

Gong, C. & Frangopol, D. M. (2020). Condition-Based Multiobjective Maintenance Decision Making for Highway Bridges Considering Risk Perceptions. *Journal of Structural Engineering,* 146(5).

Grether, D. M. (1980). Bayes Rule as a Descriptive Model: The Representativeness Heuristic. *Quarterly Journal of Economics,* 95(3), 537-557.

Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: some experimental evidence. *Journal of Economic Behavior and Organization,* 17(1), 31-57.

Griffin, D. & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology,* 24(3), 411-435.

Griffiths, T. L. & Tenenbaum, J. B. (2007). From mere coincidences to meaningful discoveries. *Cognition,* 103(2), 180-226.

Kahneman, D. & Tversky, A. (1972). Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology,* 3, 430-454.

Kahneman, D. & Tversky, A. (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology,* 5(2), 207-232.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological review,* 80(4), 237-251.

Kahneman, D. & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica,* 47(2), 263-292.

Melchers, R. E. (1999). *Structural reliability: analysis and prediction.* 2nd ed. Chichester: John Wiley & Sons.

Neumann, J. V. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Parmigiani, G. & Inoue, L. (2009). *Decision Theory: Principles and Approaches.* Chichester: Wiley.

Raiffa, H. & Schlaifer, R. (1961). *Applied Statistical Decision Theory.* Boston: Clinton Press.

Romanycia, M. H. J. & Pelletier, F. J. (1985). What is a heuristic? *Computational Intelligence*, 1(1), 47-58.

Sivia, D. & Skilling, J. (2006). *Data analysis: A Bayesian Tutorial.* Oxford: Oxford University Press.

Tenenbaum, J. B. & Griffiths, T. (2001). *The rational Basis of Representativeness.* Proc. 23rd Annual Conf. of the Cognitive Science Society, Edinburgh.

Tonelli, D., Verzobio, A., Bolognani, D., Cappello, C., Glisic, B., Zonta, D. & Quigley, J. (2018). *The conditional value of information of SHM: what if the manager is not the owner?*. Proceedings of SPIE, Denver (USA).

Tonge, F. M. (1960). Summary of a heuristic line balancing procedure. *Management Science,* 7(1), 21-42.

Tversky, A. & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, New Series*, 185, 1124-1131.

Tversky, A. & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review,* 90(4), 93-315.

Verzobio, A., Bolognani, D., Zonta, D. & Quigley, J. (2019). *Quantifying the benefit of structural health monitoring: can the value of information be negative?.* Proceedings of the 12th International Workshop on Structural Health Monitoring, Stanford (USA).

Zonta, D., Glisic, B. & Adriaenssens, S. (2014). Value of information: impact of monitoring on decision-making. *Structural Control Health Monitoring,* 21, 1043-1056.

Zonta, D., Zandonini, R. & Bortot, F. (2007). A reliability-based bridge management concept. *Structure and Infrastructure Engineering,* 3(3), 215-235.

Table 1. Achieved results for each model.

	Likelihood $P(C_i \mid h_j)$ or Representativeness $R(C_i \mid h_j)$	Posterior probability $P(h_j \mid C_i)$	Posterior odds $P(h_j \mid C_i)/P(-h_j \mid C_i)$
Bayes	$P(C_1 \mid h_{FC}) = 0.0313$ $P(C_2 \mid h_{FC}) = 0.0313$ $P(C_1 \mid h_{MH}) = 0.4437$ $P(C_2 \mid h_{MH}) = 0.0024$	$P(h_{FC} \mid C_1) = 57.27\%$ $P(h_{MH} \mid C_1) = 42.73\%$	$\frac{P(h_{FC} \mid C_1)}{P(h_{MH} \mid C_1)} = 1.34$
Vanilla model (Equation (4) and (7))	$R(C_1 \mid h_{FC}) = 0.07$ $R(C_2 \mid h_{FC}) = 13.04$ $R(C_1 \mid h_{MH}) = 14.18$ $R(C_2 \mid h_{MH}) = 0.08$	$P(h_{FC} \mid C_1)^{st} = 0.49\%$ $P(h_{MH} \mid C_1)^{st} = 99.51\%$	$\frac{P(h_{FC} \mid C_1)^{st}}{P(h_{MH} \mid C_1)^{st}} = 0.01$
Bordalo et al. (θ=0.8)	$R(C_1 \mid h_{FC}) = 0.07$ $R(C_2 \mid h_{FC}) = 13.04$ $R(C_1 \mid h_{MH}) = 14.18$ $R(C_2 \mid h_{MH}) = 0.08$	$P(h_{FC} \mid C_1)^{st} = 1.88\%$ $P(h_{MH} \mid C_1)^{st} = 98.12\%$	$\frac{P(h_{FC} \mid C_1)^{st}}{P(h_{MH} \mid C_1)^{st}} = 0.02$
Grether (α=0; β_1=0.8; β_2=0.2)	$R(C_1 \mid h_{FC}) = -2.65$ $R(C_2 \mid h_{FC}) = 2.57$ $R(C_1 \mid h_{MH}) = 2.65$ $R(C_2 \mid h_{MH}) = -2.57$	$P(h_{FC} \mid C_1)^{st} = 18.05\%$ $P(h_{MH} \mid C_1)^{st} = 81.95\%$	$\frac{P(h_{FC} \mid C_1)^{st}}{P(h_{MH} \mid C_1)^{st}} = 0.22$

Table 2. SE #12 concrete slab: state description for each Condition State (CS).

CS	State description of the slab surface
1	No delamination, spalling or water infiltration.
2	Possible delamination, spalling or water infiltration. Possible segregation and consequently reinforcement exposure.
3	Previously repaired or subjected to delamination or spalling. Segregation and consequently reinforcement exposure. Limited water infiltration.
4	Extended parts previously repaired or subject to delamination or spalling; deep segregation phenomena with extended exposure of reinforcement. Extended water infiltration.
5	Deep deterioration or anomalies. Reinforcement corrosion and cross-section loss require a deep analysis to verify the structural safety of the element.

Table 3. Achieved results for each model.

Model	Likelihood $P(C_i \mid h_j)$ or Representativeness $R(C_i \mid h_j)$	Posterior probability $P(h_j \mid C_i)$	Posterior odds $P(h_j \mid C_i)/P(-h_j \mid C_i)$
Bayes	$P(CS_5 \mid h_F) = 0.98$ $P(CS_5 \mid h_S) = 0.05$	$P(h_F \mid CS_5) = 1.92\%$ $P(h_S \mid CS_5) = 98.08\%$	$\frac{P(h_F \mid CS_5)}{P(h_S \mid CS_5)} = 0.02$
Vanilla model (Equation (4) and (7))	$R(CS_5 \mid h_F) = 19.6$ $R(CS_5 \mid h_S) = 0.05$	$P(h_F \mid CS_5)^{st} = 99.75\%$ $P(h_S \mid CS_5)^{st} = 0.25\%$	$\frac{P(h_F \mid CS_5)^{st}}{P(h_S \mid CS_5)^{st}} = 399$
Bordalo et al. (θ=0.8)	$R(CS_5 \mid h_F) = 19.6$ $R(CS_5 \mid h_S) = 0.05$	$P(h_F \mid CS_5)^{st} = 69.97\%$ $P(h_S \mid CS_5)^{st} = 30.03\%$	$\frac{P(h_F \mid CS_5)^{st}}{P(h_S \mid CS_5)^{st}} = 2.33$
Grether (α=0; β_1=0.8; β_2=0.2)	$R(CS_5 \mid h_F) = 2.98$ $R(CS_5 \mid h_S) = -2.98$	$P(h_F \mid CS_5)^{st} = 73.19\%$ $P(h_S \mid CS_5)^{st} = 26.81\%$	$\frac{P(h_F \mid CS_5)^{st}}{P(h_S \mid CS_5)^{st}} = 2.73$

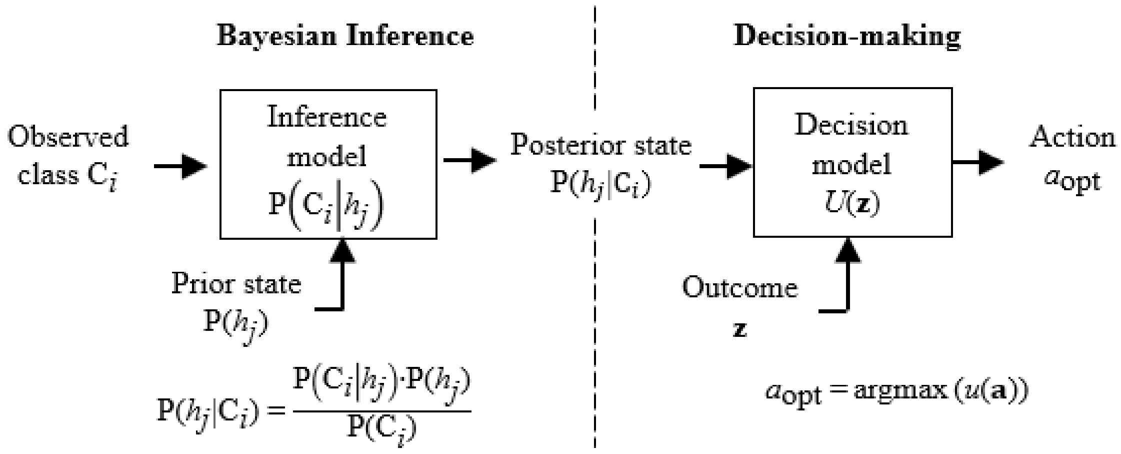
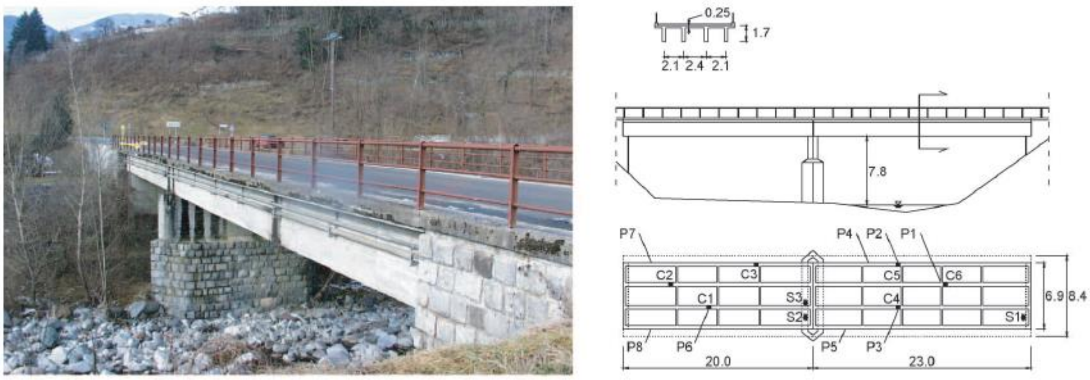


Figure 1. The rational process of SHM-based decision-making.



(a) (b)

Figure 2. SP65 bridge on Maso River: (a) overview; (b) plan view, elevation and cross section of the deck (Zonta, et al., 2007).

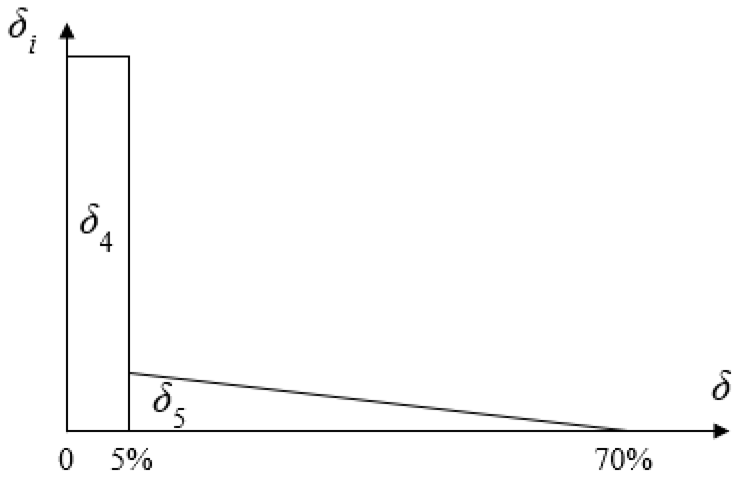


Figure 3. Capacity degradation function $\delta(CS_i)$.

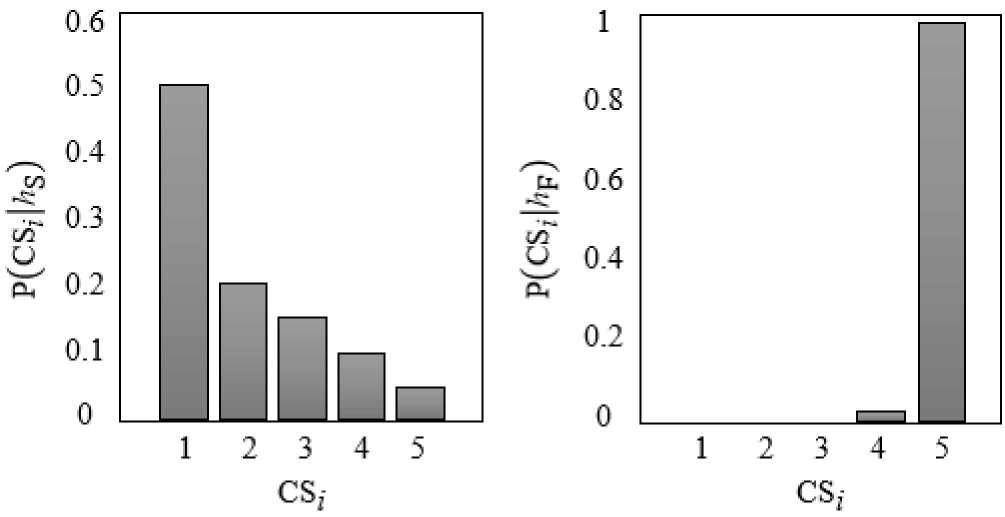


Figure 4. Likelihood distributions for each state hypothesis.

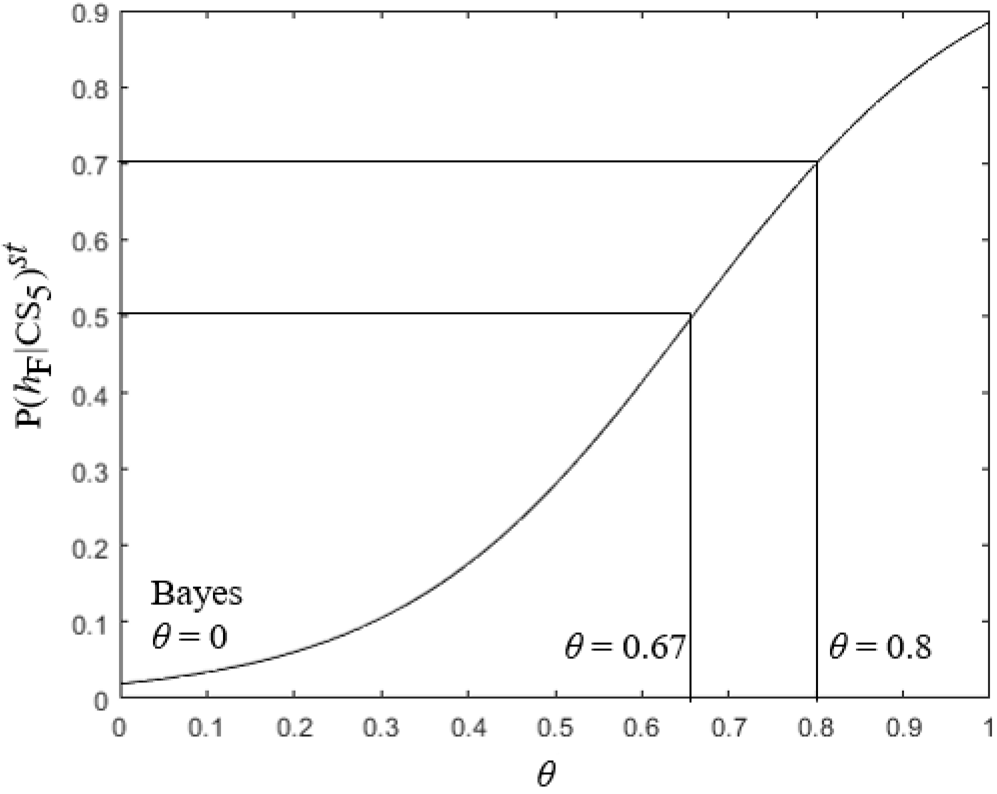


Figure 5. How the distorted posterior probability $P(h_F|CS_5)^{st}$ varies according to the subjective parameter θ.