Citation (APA)
Sun, G., Kroesen, M., & Rezaei, J. (2026). Anchoring Bias in the Tradeoff Procedure Within Multi-Attribute Value Theory. *Journal of Behavioral Decision Making*, *39*(2), Article e70069. https://doi.org/10.1002/bdm.70069

Journal of Behavioral Decision Making

WILEY

RESEARCH ARTICLE OPEN ACCESS

Anchoring Bias in the Tradeoff Procedure Within Multi-Attribute Value Theory

Geqie Sun | Maarten Kroesen | Jafar Rezaei

Faculty of Technology, Policy and Management, Delft University of Technology, Delft, the Netherlands

Correspondence: Geqie Sun (g.sun-3@tudelft.nl)

Received: 7 July 2025 | **Revised:** 21 January 2026 | **Accepted:** 23 January 2026

Keywords: anchoring bias | best–worst tradeoff | debiasing | tradeoff | weight elicitation

ABSTRACT

Eliciting the weights of attributes is a key step in multi-attribute decision-making methods. The weights usually represent the relative importance of the attributes or the tradeoffs among them in forming a decision. Various weight elicitation methods exist, each based on different assumptions and procedures. Still, many of these methods do not explicitly account for the potential influence of cognitive biases in their design. This study examines the anchoring bias, a well-known cognitive bias, in the weight elicitation step (the Tradeoff procedure) of multi-attribute value theory (MAVT). We developed the following three hypotheses: (i) Using the most important (best) attribute to construct the indifference pairs in the Tradeoff procedure leads to higher weights for the best and worst attributes and lower weights for the other attributes, (ii) using the least important (worst) attribute to construct the indifference pairs in the Tradeoff procedure leads to lower weights for the best and worst attributes and higher weights for the other attributes, and (iii) using both best and worst attributes to construct the indifference pairs (i.e., the best–worst tradeoff: BWT) mitigates the anchoring bias. To test the hypotheses, we conducted an experiment by designing a questionnaire based on MAVT and collected data from 336 participants for a decision problem. The findings indicate that the anchoring bias has a significant impact on the Tradeoff procedure and that the BWT is effective in mitigating this bias.

1 | Introduction

Many real-world decision problems are multi-attribute decision-making (MADM) problems, ranging from everyday life decisions to corporate or national governance problems, and in different fields such as healthcare, engineering, and finance. In MADM, the decision-maker (DM) evaluates multiple, often conflicting objectives. These objectives are represented by attributes, which provide measurable scales against which alternatives can be evaluated. Numerous MADM methods have been developed to support this process and to improve decision quality. One of the most important steps common to these methods is the elicitation of attribute weights. The weights usually represent the relative importance, contribution, or tradeoff among the attributes in forming a decision. These weights influence the final decision through various mechanisms, such as explicit aggregation, pairwise comparisons, or threshold-based rules. Consequently, if the weight elicitation process introduces biases or fails to capture the DM's preferences appropriately, the decision results may be misaligned with the DM's preferences. Therefore, ensuring that the elicited weights meaningfully reflect the DM's preferences is fundamental in supporting decision quality.

There are various methods for eliciting attribute weights, such as the Tradeoff procedure (Keeney and Raiffa 1976), simple multi-attribute rating technique (SMART) (Edwards 1977), Swing (von Winterfeldt and Edwards 1993), analytic hierarchy process (AHP) (Saaty 1980), and best–worst method (BWM) (Rezaei 2015), among others. Each method has distinct assumptions and procedures to help the DM elicit their preferences and translate them into quantifiable weights.

Consequently, different MADM problems can be solved by different methods depending on the context of the decision (Belton and Stewart 2012). Importantly, the interpretation of attribute weights also varies across methods (Hämäläinen and Salo 1997). As noted by Choo et al. (1999), in methods based on multi-attribute value theory (MAVT) (Keeney and Raiffa 1976), such as the Tradeoff procedure, weights have a compensatory meaning: Lower performance in one attribute can be offset by higher performance in another, depending on the weights. In contrast, methods such as AHP derive weights from pairwise comparisons and interpret weights as marginal contributions of each attribute to the overall value of an alternative, typically assuming a ratio scale. Understanding these differences is crucial, as it allows analysts to pose appropriate questions to the DM to elicit the relevant weights in different weight elicitation procedures.

These methods usually rely on the assumption that the DM is rational. In practice, however, humans are subject to cognitive biases that can systematically distort their judgments. As a result, even when these methods are implemented correctly, the elicited outputs may deviate from the normative expectations these methods are designed to reflect.

Cognitive bias is a systematic pattern of deviation from rationality when people process information. It arises from the use of mental shortcuts, or heuristics, in the decision-making process (Tversky and Kahneman 1974; Gilovich et al. 2002). While heuristics can help the DM to arrive at a good enough decision quickly, they can also lead to systematic errors in judgments and decision-making. Research on cognitive biases has therefore been largely descriptive, seeking to identify new types of bias and to develop theories or models that explain these systematic deviations from rationality (Montibeller and von Winterfeldt 2024). For instance, the anchoring bias is one of the most documented cognitive biases and has been identified in various areas (Tversky and Kahneman 1974; Chapman and Johnson 1999; Mussweiler and Strack 2001). In the medical context, Ly et al. (2023) demonstrated that prereviewing the document of a patient in the emergency department creates an anchoring bias in physicians' clinical decisions. In making a diagnosis, physicians can rely heavily on documentation and medical history (anchors), overlooking the possibility of other health conditions, leading to delays in proper examinations and diagnosis.

The insights from this descriptive line of research are highly relevant for MADM. They suggest that the fundamental assumptions (i.e., a rational DM) of MADM methods might not hold in practice, undermining the reliability of MADM methods. Consequently, research on cognitive bias in MADM has been primarily prescriptive, with an emphasis on developing debiasing strategies to account for these biases. The purpose of a prescriptive study, such as the present one, is not just to observe the biases but also to improve the decision-making process by identifying and mitigating systematic deviations from normative decision rules. In doing so, prescriptive research provides actionable insights for MADM practitioners and contributes to improving the reliability of MADM methods.

Montibeller and von Winterfeldt (2015) and Morton and Fasolo (2009) have conducted reviews of the possible cognitive biases within the context of MADM and suggested mitigation strategies. Experimental studies have also been conducted to provide empirical evidence of cognitive bias or behavioral influences in MADM methods (von Nitzsch and Weber 1993; Fischer et al. 1987; Fischer 1995; Weber et al. 1988; Weber and Borcherding 1993; Pöyhönen and Hämäläinen 2000; Rezaei 2021; Rezaei et al. 2022, 2024; Sun et al. 2025). Sun et al. (2025) investigated the anchoring bias in the value function elicitation step of MAVT, focusing on the midvalue splitting procedure. The present study complements and extends that work by shifting attention to the weight elicitation step, specifically the Tradeoff procedure. By examining another key stage of MAVT, this study deepens our understanding of how cognitive biases can arise at multiple stages of MADM methods, underscoring both the pervasiveness of such biases and the need for comprehensive debiasing strategies. Rezaei et al. (2024) examined the anchoring bias in the SMART and Swing methods. In those methods, the bias arises from explicit starting points, and their findings suggest that the BWM, which incorporates two directional anchors, can mitigate this effect. By contrast, the anchoring bias in the Tradeoff procedure does not stem from a starting point but from the selection of the reference attribute. This difference highlights that even within the same weight elicitation step of MAVT, distinct procedures can introduce bias through different mechanisms, requiring procedure-specific debiasing strategies. This study advances the literature by (i) theorizing and empirically demonstrating the anchoring bias in the Tradeoff procedure itself and (ii) testing the best–worst tradeoff (BWT) as a prescriptive debiasing strategy. While Liang et al. (2022) proposed that BWT might mitigate the anchoring bias in the Tradeoff procedure, empirical evidence has so far been lacking. By providing such evidence, the present study contributes to both the theoretical understanding of how cognitive bias influences different weight elicitation methods and the practical reliability of MADM applications.

This study investigates the effect of the anchoring bias in the weight elicitation step (the Tradeoff procedure) of MAVT. The Tradeoff procedure involves using attributes (best or worst) to construct indifference pairs and then deriving the weights (or scaling constants) from the indifference relations. We hypothesize that the selection of attributes (best or worst) might distort individual judgments due to the anchoring bias and lead to inconsistent weights across different conditions. We also hypothesize that the BWT method that utilizes both best and worst attributes to form the indifference pairs can mitigate the anchoring bias in the Tradeoff procedure. To test these hypotheses, we designed a questionnaire based on the MAVT methodology. This questionnaire contained two decision problems: The first, with two attributes, was designed for the study reported in Sun et al. (2025) on value function elicitation; the second, with three attributes, was designed for the present study to investigate weight elicitation. Although both used the apartment selection context, the attribute ranges and experimental designs were different, ensuring independent datasets despite being combined into a single instrument for efficiency. For the current study, we employed a within-subject design to compare different Tradeoff procedures. We collected data from 336 participants and performed statistical analysis to test the hypotheses.

The remainder of this paper is structured as follows: Section 2 introduces the anchoring bias and its effects in MADM methods, especially in the weight elicitation step. Section 3 describes MAVT, the Tradeoff procedure, and the BWT. Section 4 develops three hypotheses for this study. Section 5 describes the experiment design. Section 6 presents the results and discussion of the study, and Section 7 concludes this study.

2 | The Anchoring Bias and Its Role in MADM

The anchoring bias was first introduced by Tversky and Kahneman (1974), explained by the anchoring-and-adjustment heuristic. According to this heuristic, individuals begin their judgment process with an anchor and then make adjustments from that anchor to arrive at a final judgment. However, the adjustments are insufficient, resulting in a final judgment that remains biased towards the anchor. This has been demonstrated in a range of experimental settings, either through estimation tasks (measured by deviation from the true value) or valuation tasks (measured by deviation from preference coherence and consistency). For example, in an estimation task, McElroy and Dowd (2007) asked participants to estimate the length of the Mississippi River after presenting them with either a low anchor (e.g., "more than 200 miles") or a high anchor (e.g., "less than 20,000 miles"). Participants' estimates were significantly biased in the direction of the anchor, with participants in the low anchor group providing significantly lower estimates (Mean = 698.5) than those in the high anchor group (Mean = 10,021.26).

Beyond the empirical evidence, the anchoring bias has been found to be impactful across a wide range of decision fields (Furnham and Boo 2011). Accordingly, most field evidence concerns valuation tasks, in which anchoring is identified through systematic distortions in preference coherence and consistency rather than deviations from an objectively correct value. In legal settings, the anchoring bias can influence sentencing decisions made by legal professionals even when the anchor is the result of a dice roll (Englich et al. 2006), though factors such as legal expertise and the relevance of the anchor can moderate this effect (Bystranowski et al. 2021). In healthcare, the anchoring bias can influence diagnoses, where a physician's initial hypothesis anchors subsequent evaluations, even when contradictory evidence emerges (Thirsk et al. 2022). In forecasting, Campbell and Sharpe (2009) found that professional economic forecasters rely heavily on recently realized values, leading to predictable forecast errors.

To reduce the effect of the anchoring bias, various mitigation strategies have been developed, which can generally be grouped into three categories:

Encouraging critical thinking and reflection on the anchor's validity (Chapman and Johnson 1999; Epley and Gilovich 2005). This category includes strategies that encourage individuals to reflect on whether the anchor is appropriate. For instance, using incentives for accuracy can enhance individuals' motivation to make careful judgments, leading to more effortful cognitive processing and greater adjustments away from the anchor. For example, Epley and Gilovich (2005) demonstrated that when individuals generated their own anchors (i.e., self-generated anchors) and were provided with financial incentives for accuracy before the task, they then made greater adjustments away from the anchor and more accurate estimates compared to those who received no incentives. This suggests that incentives enhance motivation to engage in more effortful thinking, which reduces the influence of the anchoring bias.

Educating participants about the anchoring bias (Adame 2016; Meub and Proeger 2016; Morewedge et al. 2015). This approach involves increasing individuals' awareness of the anchoring bias, either implicitly through well-designed tasks or by explicitly providing information about the bias. The goal is to help individuals recognize how anchoring may influence their judgments. Meub and Proeger (2016) found that the anchoring bias is reduced through repeated forecasting tasks, suggesting that participants learned from experience and adapted their judgment strategies over time, becoming less influenced by the anchor in the later stages of the experiment.

Structuring decision-making processes to minimize reliance on arbitrary anchors (Mussweiler et al. 2000). This category focuses on altering the structure of the task to reduce the anchor's influence. A well-known strategy in this group is the "consider-the-opposite" strategy, which encourages individuals to critically evaluate their initial judgments by considering contradictory information or alternative scenarios, thereby reducing the influence of the anchor. Mussweiler et al. (2000) demonstrated the effectiveness of this strategy through two experiments, showing that asking participants to consider why an anchor might be inappropriate can reduce the anchoring bias.

The first two categories aim to shape how individuals think about or respond to anchors, and these strategies should be conducted before the actual judgment task. However, empirical studies have found that their effectiveness in reducing the anchoring bias is often limited (Chapman and Johnson 2002; Wilson et al. 1996; Epley and Gilovich 2005). The third category aims to directly alter the structure of the decision-making context and has shown more consistent success in mitigating the anchoring bias (Nagtegaal et al. 2020; Mussweiler et al. 2000). This structural approach resonates with the logic of MADM methods, which aim to help the DM make informed decisions by providing a structured way to evaluate the attributes and alternatives. While these methods are prescriptive in nature—intended to improve decision quality—they are typically developed under the assumption that the inputs, such as the weights, are derived from consistent human judgments. As a result, the design of many MADM methods does not explicitly account for cognitive biases, leaving them susceptible to a wide range of cognitive biases, such as the anchoring bias.

Several studies have examined the anchoring bias in MADM methods. In MADM, elicited judgments constitute valuation tasks, since they reflect subjective tradeoffs and preferences rather than estimates of objectively verifiable values. Therefore, most empirical studies define the anchoring bias in MADM as systematic deviations from preference coherence and consistency in the elicited judgments or resulting decision outcomes. Sun et al. (2025) examined the anchoring bias in MAVT during the value function elicitation step. They found that the analyst's choice of starting point in the midvalue splitting procedure can systematically bias judgments, lead to biased value functions, and consequently biased decision results. Their study also tested "counter-anchor" and "no anchor" debiasing strategies, both of which effectively reduced this bias.

Buchanan and Corner (1997) investigated the effect of the anchoring bias in a multi-objective production scheduling decision problem with two interactive methods: the Zionts and Wallenius method (Zionts and Wallenius 1983) and the Free Search method (Buchanan 1997). The Zionts and Wallenius method starts from the current solution, and the DM iteratively chooses preferred directions of improvement until no further direction is preferred. In contrast, the Free Search method lets the DM explore the solution space directly, without a fixed starting point. The results showed that the decision outcome was significantly affected by the fixed starting point in the Zionts and Wallenius method, whereas no such effect was observed in the Free Search method. This highlights how structures designed to aid the decision-making process can themselves become a source of bias and unintentionally introduce or amplify the anchoring bias.

Different MADM methods have distinct procedures, and research shows that their structural design plays a crucial role in determining how susceptible the methods are to the anchoring bias. Methods that rely on a single directional anchor, such as SMART and Swing, have been shown to be particularly vulnerable. In SMART, scoring begins from the least important attribute, whereas in Swing, it begins from the most important. Despite this difference, both methods show the same pattern of the anchoring bias in an estimation task: The weights for the less important attributes are higher than their actual ones, while the weights for the more important attributes are lower than their actual ones (Rezaei 2021). Building on this work, Rezaei et al. (2024) analyzed BWM, which incorporates two directional anchors through pairwise comparisons: one between the best attribute and the remaining attributes and another between the worst attribute and the remaining attributes. The results showed that, compared to SMART and Swing, which rely on single-directional anchors, BWM can produce lower weights for the less important attributes and higher weights for the more important attributes. In another study, Rezaei (2022) demonstrated that methods relying on a single directional anchor are more prone to the anchoring bias, while the two-directional structure in BWM reduces its impact. These findings suggest that MADM methods with multiple or opposite anchors are less susceptible to the anchoring bias.

In summary, these studies highlight the significant role of method structure in shaping decision outcomes. While existing MADM methods like SMART and Swing are thoughtfully designed and account for many theoretical and practical considerations, they were not originally developed with the impact of cognitive biases in mind. Insights from behavioral decision research have since shown that preferences are not merely revealed but are often constructed during the elicitation process and are context-dependent (Slovic 1995; Payne et al. 1992). This perspective reinforces the prescriptive function of MADM methods, which aim to guide the DM toward more informed decisions through structured procedures. On the one hand, if the design of these procedures does not account for cognitive biases from the outset, these procedures might become the source of the cognitive biases and fail the prescriptive aim of improving decision quality. On the other hand, recognizing that preferences are constructed during the elicitation process also opens the door to effective debiasing, as long as these behavioral aspects are considered in the method structure.

As one of the most widely used weight elicitation methods, the Tradeoff procedure has been included in prior investigations of behavioral effects. For instance, Weber and Borcherding (1993) examined several cognitive influences, such as range sensitivity, splitting bias, hierarchical structuring, and framing effects, across different elicitation methods, including the Tradeoff procedure. However, their work did not address the potential for the anchoring bias within the Tradeoff procedure's structure. This study focuses specifically on the anchoring bias, one of the most well-known cognitive biases (Tversky and Kahneman 1974; Furnham and Boo 2011), in the Tradeoff procedure, thereby extending this line of research by identifying and examining a previously overlooked behavioral vulnerability. The present contribution should be differentiated from recent studies by Rezaei and colleagues, who examined the anchoring bias in SMART and Swing. In those methods, anchoring arises from an explicit anchor, whereas in the Tradeoff procedure, the bias originates from the choice of reference attribute. This broadens our understanding of how biases can enter MADM methods, and practically, it calls for method-specific strategies rather than general ones.

3 | An Overview of the Tradeoff Procedure

To introduce the Tradeoff procedure, it is essential to first position it within the context of multi-attribute value theory (MAVT). MAVT is a widely recognized MADM method developed by Keeney and Raiffa (1976). It assumes that a DM's preferences can be represented by a value function consisting of attribute-specific value functions and corresponding weights (scaling constants). The Tradeoff procedure was originally developed within this theory to elicit the scaling constants by identifying indifference points between hypothetical alternatives. Since the calculation of weights using the Tradeoff procedure relies on the existence and structure of the underlying value function, MAVT provides the necessary conceptual foundation. This section introduces MAVT, the Tradeoff procedure, and the BWT.

3.1 | MAVT

The first step of MAVT involves clearly defining the decision-making context. This includes identifying the objectives, attributes, and alternatives. Let $\mathscr{A} = \{a_1, a_2, \ldots, a_M\}$ denote the set of alternatives being considered. Each alternative $a_i \in \mathscr{A}$ is evaluated based on a set of attributes $\mathscr{X} = \{X_1, X_2, \ldots, X_N\}$, where each attribute $X_j \in \mathscr{X}$ is used to assess how well an alternative meets the objectives. The goal is to determine which alternative best satisfies the objectives based on the evaluation of the attributes.

Once the decision-making context is established, the next step is to elicit attribute-specific value functions. An attribute-specific value function represents how a DM values different levels of performance for a specific attribute on a scale from 0 (the least preferred level) to 1 (the most preferred level). There are different value function elicitation procedures such as the midvalue splitting procedure, the lock-step procedure, the standard difference procedure, successive comparison, and curve fitting, among others (Keeney and Raiffa 1976; Beinat 1997; Fishburn 1967; Watson and Buede 1987). The midvalue splitting procedure is used in the current study because it is one of the most commonly employed methods originally introduced by Keeney and Raiffa (1976). In the midvalue splitting procedure, the DM identifies several midvalue points within each attribute's range, which are used to plot the value function. When a higher attribute value is preferred, an increasing function is applied. A formal definition of the midvalue is (Kirkwood and Sarin 1980) as follows:

Definition 1. $x_j^{0.5}$ is said to be the midvalue of the interval $[x_j^0, x_j^1]$ if the DM will give up the same amount of some other attribute to go from x_j^0 to $x_j^{0.5}$ as from $x_j^{0.5}$ to x_j^1.

A similar definition can be used for the decreasing attribute-specific value function.

The third step involves eliciting the attribute weights (or scaling constants). Different weight elicitation methods can be used, such as SMART (Edwards 1977) and Swing (von Winterfeldt and Edwards 1993). The Tradeoff procedure is considered in this study because it is the one developed originally in MAVT, which will be introduced in detail in the next subsection.

Once the attribute-specific value functions and weights have been elicited, the next step is to aggregate them into an overall value for each alternative. Different aggregation models can be applied depending on the underlying preference assumptions (Keeney 1974). The additive model is the simplest and most widely used. It is valid under the assumptions of mutual preferential independence and difference independence (Smith and Dyer 2021; Keeney and Raiffa 1976).

Definition 2. Attributes $X_1, \ldots, X_N$ are mutually preferentially independent if every subset of attributes is preferentially independent of its remaining attributes.

Definition 3. Attribute X_j is difference independent of the remaining attributes if the preference difference between any two levels of X_j is not affected by the fixed levels on the other attributes.

Under these assumptions, the additive model is expressed as follows:

$$v(a_i) = \sum_{j=1}^{N} w_j v_j(a_{ij}) \quad (1)$$

Here, $v(a_i)$ represents the overall value of alternative a_i, scaled from 0 to 1. $v_j(a_{ij})$ denotes the attribute-specific value of alternative a_i with respect to attribute X_j, and w_j is the scaling constant (or weight) associated with attribute X_j.

The final step of MAVT is to rank the alternatives based on their overall values or select the best-performing alternative. Specifically, for any two alternatives a_k and a_l,

$$v(a_k) \geq v(a_l) \Leftrightarrow a_k \succsim a_l$$

where the symbol $\succsim$ denotes "preferred or indifferent to" (Keeney and Raiffa 1976) in the sense of the ordering implied by the overall value function (i.e., the additive model).

In MAVT, the overall value of an alternative is computed as a weighted sum of attribute-specific value functions if the preferential independence conditions hold, as shown in the aggregation model (1). This model implies that the outcome depends jointly on two elements: the attribute-specific value functions and the weights (scaling constants). If either component is biased, the overall evaluation will also be distorted. Thus, even if the attribute-specific value functions are properly elicited, the results can only be meaningful if the weights are also properly specified. It is, therefore, essential to ensure that the weights elicited reflect the DM's preference rather than being distorted by the anchoring bias during the Tradeoff procedure.

3.2 | Tradeoff Procedure

The Tradeoff procedure typically involves the following steps:

Define attribute ranges. To define the full range of each attribute, it is first necessary to identify all possible performance levels that the attribute can take in the given decision context, within the compensatory range. Let $\underline{x}_j$ and $\overline{x}_j$ denote the worst and best performance levels for attribute X_j, respectively, where $j \in \{1, 2, \ldots, N\}$ for N attributes. Thus, $v_j(\underline{x}_j) = 0$ and $v_j(\overline{x}_j) = 1$, for all j.

Identify the importance order. The DM is presented with a set of hypothetical alternatives, each representing an extreme combination of attribute values. For a set of n attributes, each alternative corresponds to a combination where one attribute is at its best performance level (denoted $\overline{x}_j$) while the others are at their worst performance levels (denoted $\underline{x}_j$). The alternatives can be represented as follows:

$$a_1 := (\overline{x}_1, \underline{x}_2, \cdots, \underline{x}_N),\ a_2 := (\underline{x}_1, \overline{x}_2, \cdots, \underline{x}_N), \cdots,\ a_N := (\underline{x}_1, \underline{x}_2, \cdots, \overline{x}_N)$$

The DM is asked to rank these alternatives based on their preferences. This ranking process reveals the importance order of the attributes. The attribute that is at its best performance level in the most preferred alternative is defined as the most important attribute for this specific problem. For instance, if the DM ranks a_1 as the most preferred alternative, then X_1 is considered as the most important attribute for the DM and will be used to construct indifference pairs in the next step.

Construct indifference pairs. The DM is presented with a series of hypothetical comparisons, each involving two alternatives, in order to establish indifference relations. In each comparison, one attribute of the first alternative is set at its best possible level, while all other attributes are fixed at their worst levels. The second alternative in each comparison involves adjusting the level of the most important attribute until the DM is indifferent between the two alternatives, with all other attributes fixed at the worst levels. The adjusted level of the most important attribute at which the DM becomes indifferent reflects the tradeoff they are willing to make between the most important attribute and the other attribute. Attributes not involved in the comparison are held constant at their worst levels in order to isolate the tradeoff between the best attribute and each of the other attributes.

This process results in $N-1$ indifference pairs for N attributes. For any indifference pair between the most important attribute X_B, or in short B, and another attribute X_k, or in short k, it can be written as follows[1]:

$$(\underline{x}_1, \ldots, \underline{x}_B, \ldots, \overline{x}_k, \ldots \underline{x}_N) \sim (\underline{x}_1, \ldots, x_B^{B,k}, \ldots, \underline{x}_k, \ldots, \underline{x}_N) \quad (2)$$

where $\sim$ indicates the indifference relation and $x_B^{B,k}$ represents the adjusted levels of the most important attribute X_B that makes the DM indifferent between the two alternatives in each indifference pair.

Calculate weights. Each indifference pair generates an equation that quantifies the tradeoff between the most important attribute and another attribute. The first alternative has the most important attribute at its worst level $\underline{x}_B$ and attribute k at its best level $\overline{x}_k$. The second alternative has the most important attribute adjusted to a level $x_B^{B,k}$, while the other remaining attributes remain at their worst levels (with a value of zero). The indifference relation implies that the total value of the two alternatives is equal. Under the additive value function specified in (1), this can be expressed by the following equation:

$$w_k \underbrace{v_k(\overline{x}_k)}_{=1} + \sum_{\substack{j=1 \\ j\neq k}}^{N} w_j \underbrace{v_j(\underline{x}_j)}_{=0} = w_B v_B(x_B^{B,k}) + \sum_{\substack{j=1 \\ j\neq B}}^{N} w_j \underbrace{v_j(\underline{x}_j)}_{=0} \quad (3)$$

which collapses into the following:

$$w_k = w_B v_B(x_B^{B,k}) \quad (4)$$

This equation reflects the tradeoff the DM is willing to make between the most important attribute B and the attribute k. For N attributes, using the $N-1$ equations obtained from the indifference pairs and the constraint that the sum of the weights is equal to one, the weights for all attributes can be calculated as follows:

$$\begin{cases} w_1 = w_B v_B(x_B^{B,1}) \\ \ldots \\ w_k = w_B v_B(x_B^{B,k}) \\ \ldots \\ w_N = w_B v_B(x_B^{B,N}) \\ \sum_{j=1}^{N} w_j = 1 \\ w_j \geq 0, \quad j = 1, 2, \ldots, N \end{cases} \quad (5)$$

The weights derived using the Tradeoff procedure are based on the DM's indifference relations between constructed indifference pairs. In this process, the DM adjusts the level of the most important attribute starting from its worst level. This initial worst level may act as an anchor, leading to insufficient adjustments when determining the level of the most important attribute required for indifference. As a result, the anchoring bias might distort the assigned scores for the most important attribute and, consequently, lead to biased weights (we discuss this in detail in Section 4).

3.3 | BWT Method

The BWT was developed by Liang et al. (2022). It builds on the traditional Tradeoff procedure by eliciting tradeoffs between the most important attribute and the remaining attributes, as well as between the least important attribute and the remaining attributes, and then using an optimization model to derive attribute scaling constants (weights) by minimizing the inconsistency in the judgments.

The BWT procedure involves six steps in total. The first two steps, which are (i) define the alternatives and attributes and (ii) determine the attribute-specific value functions, are the same as in MAVT and will not be repeated here. The remaining steps are as follows:

Identify best and worst attribute. The DM in BWT needs to identify both the best and the worst attributes. This identification process follows a similar logic to the "*Identify the Importance Order*" step in Section 3.2. The DM is presented with a set of hypothetical alternatives, each constructed such that one attribute is fixed at its best performance level, while all other attributes are set to their worst performance levels. From the DM's ranking of these alternatives, the importance order of the attributes is derived. Throughout the paper, we use the terms *best* and *worst* to indicate the most and least important attributes, respectively. Accordingly, the most important attribute is denoted as the best attribute (B) and the least important attribute as the worst attribute (W). These two attributes are then used to construct the indifference pairs in the next steps.

Best-to-Others (BTO) tradeoff. This step mirrors the traditional Tradeoff procedure, in which the best attribute is compared against each of the remaining attributes to establish tradeoff relations. For a decision problem involving N attributes, a total of $N-1$ indifference pairs are elicited between the best attribute B and every other attribute, j, as described in (2).

If the DM reaches indifference when the level of the best attribute is adjusted to $x_B^{B,j}$, the value at that point, $v_B(x_B^{B,j})$, is denoted by b_{jB}, which represents the DM's elicitation of the ratio w_j/w_B. The reciprocal, $b_{Bj} = 1/b_{jB}$, then corresponds to the ratio w_B/w_j. To facilitate the mathematical representation, the BWT method defines the BTO vector $\mathbf{b}^{BO} = \{b_{B1}, b_{B2}, \ldots, b_{BN}\}$, where each element b_{Bj} captures the tradeoff information between the best attribute B and another attribute j in the BTO comparisons. While the traditional Tradeoff procedure often constructs only the BTO indifference pairs, BWT also considers the tradeoff between the other attributes and the worst attribute.

Others-to-Worst (OTW) tradeoff. This step involves comparing the other attributes to the worst attribute. For a decision problem with N attributes, this step also involves constructing $N-1$ indifference pairs, and each pair compares an attribute j, $j \neq W$, with the worst attribute W. These indifference pairs determine how much of an improvement in X_j is needed for the DM to consider it equivalent in changes of the worst attribute from its worst level to the best level. This process results in $N-1$ indifference pairs for N attributes. For any indifference pair between the least important attribute W and another attribute k, it can be written as follows:

$$(\underline{x}_1, \ldots, x_k^{k,W}, \ldots, \underline{x}_W, \ldots, \underline{x}_N) \sim (\underline{x}_1, \ldots, \underline{x}_k, \ldots, \bar{x}_W, \ldots, \underline{x}_N) \quad (6)$$

Similar to the traditional Tradeoff procedure, the indifference relations can also be expressed by equations using the additive value function (1), which is as follows:

$$w_k v_k(x_k^{k,W}) + \sum_{\substack{j=1 \\ j\neq k}}^{N} w_j \underbrace{v_j(\underline{x}_j)}_{=0} = w_W \underbrace{v_W(\bar{x}_W)}_{=1} + \sum_{\substack{j=1 \\ j\neq W}}^{N} w_j \underbrace{v_j(\underline{x}_j)}_{=0} \quad (7)$$

which collapses into the following:

$$w_k v_k(x_k^{k,W}) = w_W \quad (8)$$

For N attributes, using the $N-1$ equations obtained from the OTW indifference pairs and the constraint that the sum of the weights is equal to 1, the weights can be calculated as follows:

$$\begin{cases} w_1 v_1(x_1^{1,W}) = w_W \\ \ldots \\ w_k v_k(x_k^{k,W}) = w_W \\ \ldots \\ w_N v_N(x_N^{N,W}) = w_W \\ \sum_{j=1}^{N} w_j = 1 \\ w_j \geq 0, \quad j = 1, 2, \ldots, N \end{cases} \quad (9)$$

Suppose the DM reaches indifference when the level of X_j is adjusted to $x_j^{j,W}$, while the worst attribute changes from its worst level to the best level. The value $v_j(x_j^{j,W})$ is denoted as b_{Wj}, representing the ratio w_W/w_j. Its reciprocal, $b_{jW} = 1/b_{Wj}$, then corresponds to the ratio w_j/w_W. The OTW vector is defined as $\mathbf{b}^{OW} = \{b_{1W}, b_{2W}, \ldots, b_{NW}\}$, where each element b_{jW} captures the OTW tradeoff between attribute j and the worst attribute W.

Find the optimal weights. After obtaining the BTO and OTW vectors, the following system of linear equations can be formed:

$$\begin{cases} b_{Bj} = \dfrac{w_B}{w_j}, \quad \forall j \neq B \\ b_{jW} = \dfrac{w_j}{w_W}, \quad \forall j \neq W \\ w_1 + w_2 + \cdots + w_N = 1 \\ w_j \geq 0, \quad j = 1, 2, \ldots, N \end{cases} \quad (10)$$

However, the elicited judgments are often inconsistent, meaning that the equation system (10) typically does not admit an exact solution. To address this, a nonlinear optimization model can be employed to derive the optimal weights, formulated as follows:

$$\begin{aligned} &\text{minimize} && \xi \\ &\text{subject to} && \left| b_{Bj} - \frac{w_B}{w_j} \right| \leq \xi, \quad \forall j \neq B \\ &&& \left| b_{jW} - \frac{w_j}{w_W} \right| \leq \xi, \quad \forall j \neq W \\ &&& w_1 + w_2 + \cdots + w_N = 1 \\ &&& w_j \geq 0, \quad j = 1, 2, \ldots, N \end{aligned} \quad (11)$$

In this model, w_j represents the weight of attribute j, B is the best attribute, and W is the worst attribute. The objective is to minimize the maximum absolute violation of the equations (denoted by ξ). This model ensures that the tradeoffs are as consistent as possible while adhering to the constraint that the sum of the weights is equal to one. BWT also includes a consistency check for indifference pairs and weights, providing a basis for the DMs to revise their judgments in the two tradeoff procedures. A linear optimization form is also presented in Liang et al. (2022).

4 | Hypotheses Development

During the tradeoff procedure, the DM is asked to compare two hypothetical alternatives that differ in their attribute levels. One alternative features a specific attribute k at its best level, while all other attributes are fixed at their worst levels. In the other alternative, all the attributes, including attribute k, are set to their worst level, and the DM adjusts the level of attribute l until the DM is indifferent between the two alternatives. This adjustment starts from an initial reference point, such as the worst level of attribute l in the second alternative, which may serve as an anchor. Due to the anchoring bias, the DM might make insufficient adjustments from this anchor, leading to distorted indifference judgments and, ultimately, biased weight elicitation.

In the BWT method, two types of tradeoff tasks are conducted: the traditional BTO tradeoffs and the OTW tradeoffs. Although the two procedures differ in direction, the potential anchoring mechanism operates similarly in both cases.

BTO tradeoffs

In the BTO tradeoffs, the DM adjusts the best attribute B from its worst level until she is indifferent between this adjustment and the full range change (from its best level to its worst level) of another attribute $k, k \neq B$. Due to the anchoring bias, the worst level of the best attribute $\underline{x}_B$ might act as an anchor for the DM and result in insufficient adjustments in defining the adjusted level $x_B^{B,k}$. That is to say, the DM might not make enough distance from the low anchor and could assign a number closer to the worst level $\underline{x}_B$ when defining $x_B^{B,k}$ due to the insufficient adjustment.

In general, for an increasing (decreasing) attribute-specific value function of attribute B, this leads to a lower (higher) adjusted level $x_B^{B,k}$, which in turn leads to a lower $v_B(x_B^{B,k})$, since $v_B(x_B^{B,k})$ is associated to the ratio w_k/w_B (See Equation (4)). If $v_B(x_B^{B,k})$ is lower due to the anchoring bias, the derived weights will be distorted accordingly. Specifically, the weight of the attributes $k, w_k, k \neq B$, may decrease while the weight of the best attribute B, w_B may increase.

OTW tradeoffs

In the OTW tradeoffs, the DM adjusts the attribute $k, k \neq W$ from its worst level until she is indifferent between this adjustment and the full range change (from its best level to its worst level) of the worst attribute W. Similar to the BTO procedure, the worst level of the attribute k, $\underline{x}_k, k \neq W$, might act as a low anchor to the DM, and result in insufficient adjustments in defining the adjusted level $x_k^{k,W}$. That is to say, the DM might assign a number closer to the worst level $\underline{x}_k$ when defining $x_k^{k,W}$ due to the insufficient adjustment.

For an increasing (decreasing) attribute-specific value function of $k, k \neq W$, this might lead to lower (higher) adjusted levels $x_k^{k,W}$ for all OTW tradeoffs, which in turn leads to a lower $v_k(x_k^{k,W})$, since $v_k(x_k^{k,W})$ represents the ratio w_W/w_k (See Equation (8)). From (8), it is easy to observe that if $v_k(x_k^{k,W})$ is lower due to the anchoring bias, the weight of the worst attribute w_W may decrease while the weight of the non-worst attribute k, $w_k, k \neq W$, may increase.

It is important to note that one tradeoff pair, the comparison between the best and worst attributes, is common to both the BTO and OTW procedures. Therefore, any differences in the resulting weights between the two procedures originate from the remaining $2(N-2)$ tradeoffs: the BTO comparisons between the best attribute and the remaining attributes (excluding the worst) and the OTW comparisons between the worst attribute and the remaining attributes (excluding the best).

Before presenting the propositions on how anchoring affects the weights in the BTO and OTW procedures, we first show how the weights are derived from each procedure separately, as follows.

4.1 | Calculating Attribute Weights Based on BTO and OTW

Assume a decision problem with N attributes $\{X_1, \ldots, X_N\}$. Let B denote the most important (best) attribute, W the least important (worst) attribute, and $\mathscr{K} := \{1, \ldots, N\} \setminus \{B, W\}$ the index set of the remaining attributes.

4.1.1 | BTO tradeoff

Eliciting the ratios $b_{Bk} = w_B/w_k$ for every $k \in \mathscr{K} \cup \{W\}$ and imposing the normalization $\sum_{j=1}^{N} w_j = 1$ gives

$$w_B = \frac{1}{1 + \sum_{j \neq B} \frac{1}{b_{Bj}}}, \qquad w_k = \frac{\frac{1}{b_{Bk}}}{1 + \sum_{j \neq B} \frac{1}{b_{Bj}}}, \quad k \neq B \tag{12}$$

4.1.2 | OTW tradeoff

If instead the judgments use the worst attribute, one elicits $b_{kW} = w_k/w_W$ for every $k \in \mathscr{K} \cup \{B\}$. Solving the same normalization condition yields the following result.

$$w_W = \frac{1}{1 + \sum\limits_{j \neq W} b_{jW}}, \qquad w_k = \frac{b_{kW}}{1 + \sum\limits_{j \neq W} b_{jW}}, \quad k \neq W \tag{13}$$

4.1.3 | Consistency without anchoring

When the tradeoff answers are internally consistent, the cross-ratio condition

$$b_{BW} = b_{Bk}\, b_{kW}, \quad \forall\, k \in \mathscr{K} \tag{14}$$

holds (see Liang et al., 2022). Under this condition, the weight vectors obtained from BTO, OTW, and BWT are identical.

4.1.4 | Effect of the anchoring bias

Anchoring typically inflates both b_{Bk} and b_{kW}, violating the above equality. Because the shared ratio $b_{BW} = w_B/w_W$ is present in *both* procedures, we treat it as fixed and attribute any differences in the resulting weights to bias-induced changes in the remaining ratios. Propositions 1 and 2 formalize how these changes push the BTO and OTW weights in opposite directions.

Proposition 1. *In the BTO procedure, the low-anchored $v_B(x_B^{B,k}), k \neq \{B, W\}$ results in an increase in w_B and w_W and a decrease in $w_k, k \neq \{B, W\}$.*

Proof. Let $Q := v_B\left(x_B^{B,k}\right)$ and $S := \sum_{j \neq \{B,k\}} v_B\left(x_B^{B,j}\right)$. From the BTO identities

$$w_B = \frac{1}{1+S+Q}, \qquad w_k = \frac{Q}{1+S+Q}, \qquad w_W = w_B\, v_B\left(x_B^{B,W}\right)$$

where $v_B\left(x_B^{B,W}\right)$ is constant in this perturbation. Differentiating,

$$\frac{\partial w_B}{\partial Q} = -\frac{1}{(1+S+Q)^2} < 0,$$
$$\frac{\partial w_k}{\partial Q} = \frac{1+S}{(1+S+Q)^2} > 0,$$
$$\frac{\partial w_W}{\partial Q} = v_B\left(x_B^{B,W}\right) \frac{\partial w_B}{\partial Q} < 0$$

Thus, a downward shift in Q (a low anchor) increases w_B and w_W but decreases w_k. □

Proposition 2. *In the OTW procedure, the low-anchored $v_k(x_k^{k,W}), k \neq \{B, W\}$ results in a decrease in w_B and w_W and an increase in $w_k, k \neq \{B, W\}$.*

Proof. Let $P := v_k\left(x_k^{k,W}\right)$ and $R := \sum_{j \neq \{W,k\}} 1/v_j\left(x_j^{j,W}\right)$. Then

$$w_W = \frac{1}{1+R+1/P}, \qquad w_k = \frac{w_W}{P}, \qquad w_B = \frac{w_W}{v_B\left(x_B^{B,W}\right)}.$$

Differentiating the following,

$$\frac{\partial w_W}{\partial P} = \frac{1}{(1+R+1/P)^2P^2} > 0, \qquad \frac{\partial w_B}{\partial P} = \frac{1}{v_B\left(x_B^{B,W}\right)}\frac{\partial w_W}{\partial P} > 0$$

$$\frac{\partial w_k}{\partial P} = -\frac{w_W}{P} + \frac{1}{P}\frac{\partial w_W}{\partial P} = \frac{1-(1+R+1/P)P^2}{(1+R+1/P)^2P^3} < 0$$

Thus, a downward shift in P decreases w_W and w_B but increases w_k, completing the proof. □

Drawing from the above discussion, we propose the following hypotheses:

Hypothesis 1: *Using the most important (best) attribute to construct the indifference pairs in the Tradeoff procedure leads to higher weights for the best and worst attributes and lower weights for the other attributes.*

Hypothesis 2: *Using the least important (worst) attribute to construct the indifference pairs in the Tradeoff procedure leads to lower weights for the best and worst attributes and higher weights for the other attributes.*

Since the BWT method utilizes both BTO and OTW tradeoffs, it inherently incorporates the "consider-the-opposite" debiasing strategy in its structure. "Consider-the-opposite" (Mussweiler et al. 2000) is one of the most effective debiasing strategies developed for reducing the anchoring bias. It encourages individuals to critically evaluate their initial judgments by considering contradictory information or alternative scenarios. In the context of MADM methods, BWT achieves this by requiring the DM to evaluate tradeoffs from two different perspectives: using the best and worst attributes to construct the indifference pairs. This is not merely a matter of increasing the number of elicited comparisons; it systematically encourages the DM to conduct tradeoffs from opposing perspectives.

As illustrated in the previous discussions, the anchoring bias tends to distort the weights in opposite directions in the BTO and OTW tradeoff procedures. Thus, when the two sets of tradeoffs are combined in the BWT method, these opposite distortions may cancel each other out, reducing the anchoring bias introduced by any single anchor. The optimization model solves for the weight that best fits both tradeoff perspectives simultaneously and leads to weights that are less likely to be biased by a single direction anchor.

Therefore, we hypothesize the following:

Hypothesis 3: *Using the Best–Worst Tradeoff method can reduce the anchoring bias in the Tradeoff procedure.*

Remark. When the decision problem involves exactly *three* attributes, the weight of the worst attribute, w_W, shows a special behavior. Consider X_1 as the best attribute, X_3 as the worst attribute, and X_2 as the remaining one. From (12) and (13), the weights for the three attributes can be obtained for BTO and OTW, respectively, as follows.

$$BTO\begin{cases} w_1 = \dfrac{1}{1+\frac{1}{b_{12}}+\frac{1}{b_{13}}} \\ w_2 = \dfrac{1}{b_{12}}\dfrac{1}{1+\frac{1}{b_{12}}+\frac{1}{b_{13}}} \\ w_3 = \dfrac{1}{b_{13}}\dfrac{1}{1+\frac{1}{b_{12}}+\frac{1}{b_{13}}} \end{cases} \qquad OTW\begin{cases} w_1 = b_{13}\dfrac{1}{1+b_{13}+b_{23}} \\ w_2 = b_{23}\dfrac{1}{1+b_{13}+b_{23}} \\ w_3 = \dfrac{1}{1+b_{13}+b_{23}} \end{cases} \tag{15}$$

If the anchoring bias does not affect the tradeoff process and the preferences elicited are fully consistent (refer to (14)), then $b_{13} = b_{12}b_{23}$ holds. Under this condition, BTO and OTW yield identical weights. Wu et al. (2024) established a theorem[2] showing that there is a unique optimal solution for the problem under a not-fully consistent comparison system when three attributes are involved. Specifically, according to Wu et al. (2024), the optimal weights are as follows:

$$\begin{cases} w_1 = \dfrac{b_{13}-\xi^*}{b_{13}+b_{23}+1} \\ w_2 = \dfrac{b_{23}+\xi^*}{b_{13}+b_{23}+1}, \text{ if } b_{12}\times b_{23} < b_{13} \\ w_3 = \dfrac{1}{b_{13}+b_{23}+1} \end{cases} \qquad \begin{cases} w_1 = \dfrac{b_{13}+\xi^*}{b_{13}+b_{23}+1} \\ w_2 = \dfrac{b_{23}-\xi^*}{b_{13}+b_{23}+1}, \text{ if } b_{12}\times b_{23} > b_{13} \\ w_3 = \dfrac{1}{b_{13}+b_{23}+1} \end{cases} \tag{16}$$

From (15) (OTW weights) and (16), it is clear that w_3 has the same weight for OTW and BWT (regardless of the consistency level of BWT).

5 | Experimental Design

To test the hypotheses developed in Section 4, we designed a hypothetical decision problem with three attributes and implemented it in an online questionnaire (Qualtrics) that followed the steps of MAVT. The questionnaire consisted of a larger experiment with two independent decision problems. The first, a two-attribute problem, has been reported elsewhere (Sun et al. 2025) and focused on the anchoring bias in value function elicitation. The current study focused on the second, a three-attribute problem specifically designed to examine the anchoring bias in weight elicitation. In this problem, participants evaluated three attributes, and the experimental variation focused on different versions of the Tradeoff procedure (BTO, OTW, and BWT).

Combining the two problems in a single questionnaire was both efficient and methodologically sound. First, the problems targeted different steps of MAVT (value functions vs. weights), so there was no conceptual overlap that could cause contamination. Second, any potential learning or familiarity effects from the first problem would apply equally across all conditions of the second problem. Since the analysis compares differences between different Tradeoff procedures (BTO, OTW, BWT), such general familiarity does not bias results in favor of one method.

The first part of the questionnaire was to inform the participants about the content of the study, including the objective, procedures, risks, and benefits. The participants were then asked to voluntarily provide their consent to participate in this study.

In the second part, participants were presented with a hypothetical decision problem. As illustrated in Table 1, an apartment selection problem was designed. The rent attribute range was set based on the current rental market conditions in the target countries of the experiment. The lower bound for commute distance was set at a sufficiently high level to ensure that the value function remains monotonic. This prevents situations where individuals might favor an intermediate commute distance over living too close to or too far from their workplace. The distance to the shopping center was deliberately given a narrow range to reduce its overall importance. This design was intended to encourage a similar attribute importance order across participants, reducing variance due to individual differences in the experiment. Importantly, while the design encouraged consistent importance rankings for the attributes, the subsequent analysis was not dependent on the specific attributes themselves but rather on their relative importance rankings as reported by each participant. That is, all analyses were conducted based on whether an attribute was ranked as most, least, or intermediately important, regardless of which physical attribute (e.g., rent, commute, shopping distance) it represented.

The third part of the experiment consisted of three tasks: eliciting attribute-specific value functions, eliciting weights, and validating the additivity assumption. While the value functions were elicited separately, we integrated the weight elicitation with the additivity assumption validation because both procedures require constructing indifference pairs (See Supporting Information Appendix A for an example of the tradeoff questions). This integration reduced the total number of indifference judgments one has to make, allows for a more efficient experimental structure, and minimizes unnecessary cognitive effort. Reducing redundancy in elicitation tasks is particularly important in multicriteria decision experiments where complex tradeoffs are involved, and quantitative judgments are required (Larichev 1992; Larichev and Brown 2000; Jaspersen and Montibeller 2015).

To implement this integration, participants first determined the importance order of the three attributes following the procedure described in Section 3. The identified best and worst attributes are then used to construct the BTO and OTW tradeoffs. Assume X_1 was the best attribute, X_3 was the worst attribute, and X_2 was the remaining attribute. In total, each participant evaluated three indifference pairs across the two tradeoff procedures (see Table 2), noting that P_2 is identical in the two tradeoff procedures.

To verify the additivity assumption, we assessed an additional set of three indifference relations. Take P_1 from the BTO tradeoffs as an example (see Table 2): Once a participant identified the indifference value $x_1^{1,2}$, we modified the original pair by changing the third attribute level from $\underline{x}_3$ to $\bar{x}_3$ and asked the participants whether they were still indifferent between the two options. If so, it indicated that the preferences between X_1 and X_2 were independent of X_3. We performed similar checks for the other two pairs, $\{X_1, X_3\}$ and $\{X_2, X_3\}$. Once all indifference pairs had been assessed, the additivity assumption was fully validated. In the formal BWT procedure, there will be necessary revisions after checking the ordinal and cardinal consistency ratios, but this step was not performed due to the experimental constraints.

Importantly, this experiment employed a within-subject design under three conditions: BTO tradeoff, OTW tradeoff, and BWT tradeoff. Since the BWT method is inherently constructed from the BTO and OTW tradeoffs, a within-subject design was particularly well-suited. By applying a within-subject design, the experiment ensured that any observed differences across the BTO, OTW, and BWT conditions can be directly attributed to the Tradeoff procedures themselves, rather than to differences among participants (Greenwald 1976). In contrast, a between-subject design would introduce interparticipant variability, such as the participant's prior experiences, knowledge, or decision-making styles, that could potentially obscure these effects (Charness et al. 2012).

Since the tradeoff procedure requires attribute-specific value functions to calculate weights, we used the midvalue splitting procedure to elicit these functions after completing the tradeoff tasks. Although this study focuses on the anchoring bias in the weight elicitation step, value functions are essential for converting indifference judgments into weights under the additive model. To ensure that any observed bias in the weights stemmed

TABLE 2 | Indifference pairs in the experiment.

Indifference pairs	Best-to-Others tradeoff	Others-to-Worst tradeoff
P_1	$(\underline{x}_1, \bar{x}_2, \underline{x}_3) \sim (x_1^{1,2}, \underline{x}_2, \underline{x}_3)$	$(\underline{x}_1, x_2^{2,3}, \underline{x}_3) \sim (\underline{x}_1, \underline{x}_2, \bar{x}_3)$
P_2	$(\underline{x}_1, \underline{x}_2, \bar{x}_3) \sim (x_1^{1,3}, \underline{x}_2, \underline{x}_3)$	$(x_1^{1,3}, \underline{x}_2, \underline{x}_3) \sim (\underline{x}_1, \underline{x}_2, \bar{x}_3)$

TABLE 1 | Attributes used in the decision problem.

Attribute	Unit	Description	Range
Monthly rent	Euro	It is the amount of money one has to pay each month to rent the apartment.	[600,1500]
Commute distance	Kilometer	This is the proximity of the apartment to your workplace.	[5,15]
Distance to the shopping center	Meter	This is the proximity of the apartment to the nearest shopping center where you buy groceries.	[100,500]

from the tradeoff elicitation itself, and not from inconsistencies in value function elicitation, all participants used the same value function elicitation procedure. Take eliciting the attribute-specific value function for the rent as an example. To obtain the first midvalue point, we asked, "*Suppose the price drop in monthly rent would be from 1500 to a certain value (r_1) or from that same value (r_1) to 600. Please assign a value to r_1 such that for the two price drops, you would accept the same increase in commute distance and shopping center.*" r_1 was identified as the first midvalue point. The subsequent midvalue points were also determined following similar questions, and the attribute-specific value function for rent could be plotted using these midvalue points (Keeney and Raiffa 1976, Section 3.4.8).

The last step was collecting demographic information such as age, gender, and education level to have a comprehensive understanding of the participant's profile. Besides, the participant's current rent or mortgage, commute distance, and distance to the shopping center were also collected as control variables.

Data were collected via the online platform Prolific (Palan and Schitter 2018), which offers various prescreening features and a response verification process that enhances data quality. We recruited a total of 440 participants from six European countries: the Netherlands, Germany, France, Belgium, Denmark, and Luxembourg. These countries share similar rental market conditions and were thus selected. Since the questionnaire was in English, we also limited participation to individuals fluent in English using the prescreening functions. Moreover, the response verification function enabled us to reject incomplete answers or those respondents that provided answers outside of the value ranges. Participants received a small monetary incentive for successful completion. Such incentives have been shown to increase response rates and improve the quality of answers (Singer and Ye 2013). Of the 440 participants recruited, 36 were excluded for not completing the experiment. An additional 68 participants were removed as they provided values outside the specified ranges or identical values across all midvalue points. This indicated either inattention or insufficient understanding of the MAVT questions. The entire questionnaire, which comprised the two decision problems, was completed on average of 16 min and 38s, with a standard deviation of 8 min and 59 s.

The final sample included 336 participants; the detailed demographics of the sample are presented in Table 3. After data collection, the statistical analyses were conducted in SPSS, Version 29. We employed both parametric and nonparametric tests. ANOVA with post hoc tests was used to capture the magnitude of the effects, while Wilcoxon signed-rank tests were applied to assess the direction of the changes. We considered both aspects important: Although the magnitude indicates the strength of the anchoring bias, the direction provides a more direct test of whether the bias systematically shifts the results in the hypothesized way.

6 | Results and Discussion

The experiment was designed to test the effect of the anchoring bias in the Tradeoff procedure within the context of MAVT. Both parametric and nonparametric tests were used to perform the analysis. Notably, ANCOVA tests were conducted to examine whether the participant's current living space, housing costs, commute distance, shopping distance, and demographics affected the main results of this study. The results indicated that none of the control variables had a statistically significant effect ($p > 0.05$ for all).

TABLE 3 | Demographic characteristics of participants ($n = 336$).

Characteristics	Levels	Percent
Gender	Male	62%
	Female	35%
	Other	3%
Age	[18,24]	25%
	[25,34]	49%
	[35,44]	15%
	> 44	11%
Education	High school	13%
	Bachelor's degree	34%
	Master's degree	33%
	Other	20%

Since the hypotheses were based on the relative importance of attributes, and participants may prioritize the three attributes differently, it was necessary to classify attributes according to each participant's individual ranking rather than by fixed attribute labels. This approach ensured that the analysis focused on differences in weights for the *best*, *worst*, and *other* attributes as perceived by each participant, rather than being influenced by the specific attribute names. Specifically, the highest ranked attribute for a participant was treated as their *best* attribute, the lowest ranked as their *worst*, and the remaining one as *other*, regardless of whether the attribute was rent, commute distance, or distance to the shopping center. In general, participants tended to prioritize rent, with 85% ranking it as their most important attribute, followed by 12% for commute distance and 3% for distance to the shopping center.

The average weights based on this classification were presented in Figure 1. The results indicated that using the Best-to-Others (BTO) tradeoffs in the tradeoff procedure resulted in higher weights for the *best*and *worst* attributes and lower weights for the *other* attribute, compared to using the OTW tradeoffs in the tradeoff procedure. The BWT method produced weights in between the two extremes for the *best* and *other* attributes and similar weights for the *worst* attribute compared to the OTW tradeoff procedure.

ANOVA tests indicated statistically significant differences in attribute weights across the three conditions: *best* attribute, $F(2, 1005) = 5.106, p = 0.006$; *other* attribute, $F(2, 1005) = 15.144, p < 0.001$; and *worst* attribute, $F(2, 1005) = 4.819, p = 0.008$. As shown in Table 4, post hoc analyses revealed that, compared to the OTW group, the BTO group

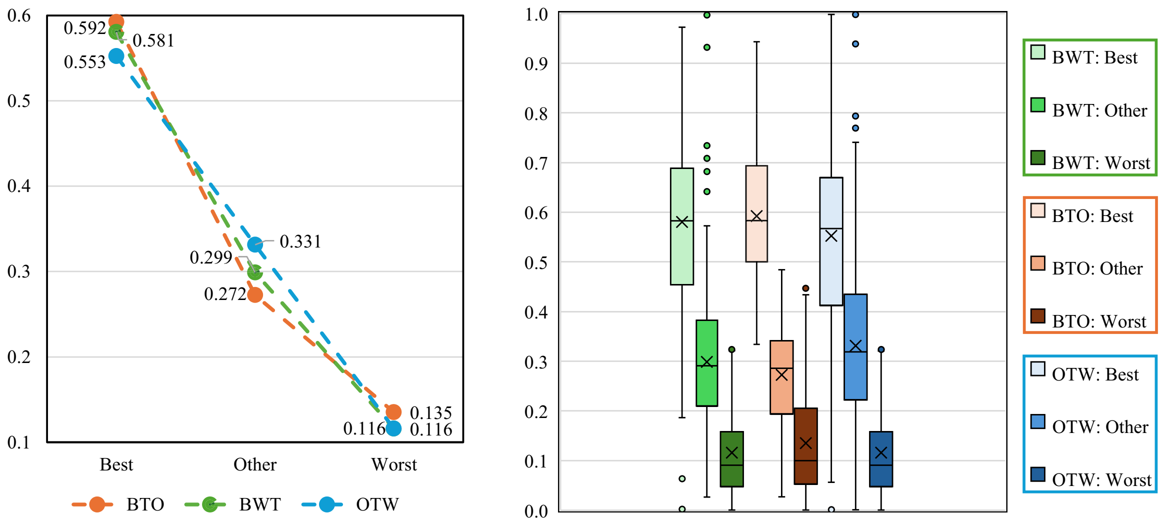

FIGURE 1 | Average weight based on individual ranking.

TABLE 4 | The mean difference of the best, other, and worst attributes of all tradeoff method pairs.

Attribute	Tradeoff method pair	Mean difference	Sig.	95% confidence interval	Maximum difference
Best	BWT-BTO	−0.012	0.359	[−0.037, 0.013]	−0.5219
Best	BWT-OTW	0.028	0.029	[−0.003, 0.053]	0.3795
Best	BTO-OTW	0.040	0.002	[0.015, 0.065]	0.5291
Other	BWT-BTO	0.027	0.014	[0.005, 0.048]	0.6543
Other	BWT-OTW	−0.033	0.002	[−0.054, −0.012]	−0.5856
Other	BTO-OTW	−0.059	< 0.001	[−0.080, −0.038]	−0.6553
Worst	BWT-BTO	−0.019	0.007	[−0.033, −0.005]	−0.3104
Worst	BWT-OTW	0.000	1.000	[−0.014, 0.014]	0.000
Worst	BTO-OTW	0.019	0.007	[0.005, 0.033]	0.3104

assigned significantly higher weights to both the *best* attribute (BTO: Mean = 0.592, OTW: Mean = 0.553, $p = 0.002$) and *worst* attribute (BTO: Mean = 0.135, OTW: Mean = 0.116, $p = 0.007$) and significantly lower weights to the *other* attribute (BTO: Mean = 0.272, OTW: Mean = 0.331, $p < 0.001$). Moreover, maximum observed differences illustrated that the anchoring bias could have a substantial impact at the individual level. For instance, the largest difference for the *other* attribute between BTO and OTW was –0.6553. These findings supported Hypotheses 1 and 2, demonstrating the presence of the anchoring bias in the Tradeoff procedure.

To test the effectiveness of BWT in producing more balanced weights and thus reducing the anchoring bias in the Tradeoff procedure, we compared the results of the BWT group with both the BTO and OTW groups. Compared to the BTO group, the BWT method resulted in significantly higher weights for the *other* attribute ($p = 0.014$) and significantly lower weights for the *worst* attribute ($p = 0.007$). Although the weights for the *best* attribute were also lower in the BWT group, this difference did not reach statistical significance ($p = 0.355$). In comparison to the OTW group, the BWT group produced significantly higher weights for the *best* attribute ($p = 0.032$) and significantly lower weights for the *other* attribute ($p = 0.002$). The BWT and OTW produced the same weights for the worst attribute ($p = 1$), consistent with expectations.

These results generally supported Hypothesis 3, indicating that the BWT method effectively reduced the anchoring bias. However, the nonsignificant difference in the best attribute between the BWT and BTO groups ($p = 0.359$) was unexpected and warrants further investigation.

While the ANOVA highlighted differences in the magnitude of weights across groups, it did not capture the direction of change. To better understand how the anchoring bias systematically shifted weight elicitation under different conditions, we conducted Wilcoxon signed-rank tests. This nonparametric analysis focused on within-subject directional changes and provided complementary evidence that could reinforce the overall pattern of anchoring effects observed in the ANOVA and post hoc analyses. As shown in Table 5, the results not only aligned with but also extended the results of the ANOVA and post hoc analyses, with the previously nonsignificant difference in the

TABLE 5 | Wilcoxon signed-rank test.

	BTOBest-BWTBest	OTWBest-BWTBest	OTWBest-BTOBest	BTOOther-BWTOther	OTWOther-BWTOther	OTWOther-BTOOther	BTOWorst-BWTWorst	OTWWorst-BWTWorst	OTWWorst-BTOWorst
Z	−2.200[b]	−5.863[a]	−5.071[a]	−4.282[a]	−6.267[b]	−5.775[b]	−7.753[b]	0.000[c]	−7.753[a]
Sig.	0.028	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	1	< 0.001

[a]Based on positive ranks.
[b]Based on negative ranks.
[c]The sum of negative ranks equals the sum of positive ranks.

best attribute between the BWT and BTO groups now reaching significance ($p = 0.028$), offering stronger support for the three hypotheses.

With respect to Hypotheses 1 and 2, participants in the OTW group assigned significantly lower weights to the *best* and *worst* attributes ($p < 0.001$ for both) and significantly higher weights to the *other* attribute ($p < 0.001$) compared to those in the BTO group. These results confirmed the presence of the anchoring bias in the Tradeoff procedure, where it distorted the weights in opposite directions.

Regarding Hypothesis 3, the Wilcoxon signed-rank test showed that the BWT group produced significantly lower weights for the *best* attribute compared to the BTO group ($p = 0.028$) and significantly higher weights than the OTW group ($p < 0.001$). For the *other* attribute, weights in the BWT group were significantly higher than those in the BTO group ($p < 0.001$) and significantly lower than those in the OTW group ($p < 0.001$). For the *worst* attribute, the BWT group assigned significantly lower weights than the BTO group ($p < 0.001$). These statistically significant results indicated that the weights produced by the BWT method consistently fell between the two extremes of the BTO and OTW procedures, except for the *worst* attribute, where BWT and OTW produced the same weights ($p = 1$). These results supported Hypothesis 3 and demonstrated that BWT was effective in mitigating the anchoring bias in the Tradeoff procedure.

To better illustrate how the anchoring bias could distort the Tradeoff procedure, we presented the results from a single participant, demonstrating how the weights elicited for the same DM varied under different conditions due to this bias.

Participant 64 identified the importance order of the three attributes as follows: rent, commute distance, and distance to shopping center. Based on the attribute ranges (see Table 1) and the indifference pairs specified for this experiment (see Table 2), three indifference judgments were elicited, as summarized in Table 6.

TABLE 6 | All indifference pairs and responses for participant 64.

Indifference pairs	Attribute comparison	Participant's answer
BTO : P_1	$(1500, 5, 500) \sim (x_1^{1,2}, 15, 500)$	$x_1^{1,2} = 1000$
OTW : P_1	$(1500, x_2^{2,3}, 500) \sim (1500, 15, 100)$	$x_2^{2,3} = 10$
BTO&OTW : P_2	$(1500, 15, 100) \sim (x_1^{1,3}, 15, 500)$	$x_1^{1,3} = 1200$

TABLE 7 | Weights for participant 64.

Tradeoff procedure	w_{Rent}	$w_{CommuteDistance}$	$w_{DistancetoShoppingCenter}$
BTO	0.600	0.250	0.150
BWT	0.593	0.264	0.143
OTW	0.571	0.286	0.143

TABLE 8 | Results of one-sample t-tests of Kendall's τ_b (test value = 1).

Tradeoff method comparison	Mean τ_b	Std. deviation	t	One-sided p
BWT-BTO	0.8077	0.0508	−37.852	< 0.001
BWT-OTW	0.8657	0.0379	−35.452	< 0.001
BTO-OTW	0.6915	0.0768	−40.171	< 0.001

After eliciting the attribute-specific value functions, the attribute weights were derived using BTO equation systems (12), OTW equation systems (13), and BWT optimization model (11). Table 7 presents the derived weights under the three tradeoff procedures.

As shown, noticeable differences emerged across the three conditions. For the most important attribute, rent, the BTO procedure yielded the highest weight (0.600), the OTW procedure yielded the lowest (0.571), and the BWT method produced an intermediate weight (0.593). For commute distance, the OTW procedure yielded the highest weight (0.286), the BTO procedure yielded the lowest (0.250), and the BWT method again produced an intermediate weight (0.264). For the least important attribute, distance to shopping center, the BTO procedure produced the highest weight (0.150), while the OTW and BWT procedures produced identical, lower weights (0.143).

Our experimental design was based on a relatively simple decision problem with three attributes. Real-world MADM applications often involve more attributes, sometimes 10 or more (Belton and Stewart 2012). When the number of attributes increases, each attribute's normalized weight tends to become smaller, which may reduce the influence of the anchoring bias in individual weights. However, the use of a simplified three-attribute setup allowed us to provide a clean and controlled test of the theoretically derived hypotheses regarding anchoring in weight elicitation, without the additional complexity that larger attribute sets might have introduced, such as cognitive overload or interaction effects (Payne et al. 1993; Keeney and Raiffa 1976). At the same time, it was important to examine whether the observed effects had practical implications for decision outcomes.

To address this, we conducted a simulation study to test whether differences across the three weight elicitation methods (BTO, OTW, and BWT) led to differences in the final decision outcomes. We constructed and evaluated a set of simulated alternatives. Specifically, we generated five random numbers between zero and one that summed to one and then mapped these numbers onto attribute levels within the experimental ranges. This procedure produced five nondominated alternatives for each simulation. Using the additive aggregation model, we combined each participant's attribute-specific value functions and weights to obtain an overall value for each alternative and thus a complete ranking of the five alternatives under each weight elicitation method. As a result, each participant produced three distinct rankings of the same set of alternatives. To assess the extent to which these rankings agreed or disagreed with each other, we applied the Kendall τ_b coefficient (Kendall 1948). This measure captured the correlation between two rankings, with a value of 1 indicating perfect agreement, 0 indicating no systematic association, and −1 indicating complete reversal of order.

To ensure that our findings were not dependent on the specific choice of alternatives, we repeated this procedure 100 times with independently generated sets of five nondominated alternatives. In each replication, we obtained τ_b values for the three method pairs (BWT-BTO, BWT-OTW, and BTO-OTW). We then calculated the average Kendall τ_b across the 100 replications, which provided a more stable estimate of the similarity in rankings produced by different weighting methods.

As shown in Table 8, the results of the one-sample t-tests showed that these average τ_b values were significantly below 1, indicating that the agreement between the rankings obtained from different weighting methods was far from perfect. This confirmed that the discrepancies in elicited weights across methods were not merely numerical differences but had practical consequences for the ranking of alternatives. Moreover, the τ_b for BTO-OTW comparison was the lowest among all, indicating that the anchoring bias pulled the two tradeoff procedures toward more extreme and divergent weights, whereas BWT helped to reduce these extremes.

These results aligned with our hypothesis, illustrating that the anchoring bias could distort the weights when only one type of tradeoff is used. Furthermore, they demonstrated that the BWT method can mitigate this bias by producing weights between the two extremes, thus reducing the influence of the anchoring bias.

7 | Conclusion

In this study, we investigated the effect of the anchoring bias in the Tradeoff procedure within MAVT. We found that the selection of the best or worst attribute to construct indifference pairs in the Tradeoff procedure can serve as an anchor and result in biased weights due to the anchoring bias. Specifically, using the best attribute leads to higher weights for the best and worst attributes and lower weights for the other attributes compared to using the worst attribute in the tradeoff procedures. Additionally, the BWT method was found to be effective in reducing the anchoring bias by incorporating both indifference pairs.

These findings align with previous research (Rezaei 2021; Rezaei et al. 2024; Rezaei 2022), indicating that weight elicitation methods using a single reference point (e.g., SMART, swing, or one vector of BWM) are susceptible to the anchoring bias, while methods employing two counter-reference points, such as the BWM, can reduce its impact. It is worth noting that although these studies all examined weight elicitation, they focused on different methods, each with distinct interpretations of weights and distinct procedures that introduce the anchoring bias. This study extends this line of research by examining the Tradeoff procedure. While Liang et al. (2022) proposed that BWT has an inherent debiasing mechanism for the anchoring bias, empirical evidence supporting this claim has been lacking. This study addresses this gap by providing empirical evidence for the debiasing effect of BWT.

This study enhances our understanding of the anchoring bias in weight elicitation methods, offering important implications for both the theoretical and practical aspects of the Tradeoff procedure. It also underscores the need to account for cognitive biases, particularly the anchoring bias, when developing and applying MADM methods. While the structure of a MADM method may unintentionally introduce such biases, thoughtful design can help mitigate their effects.

The limitations of this study should be acknowledged. First, the sample is skewed toward highly educated participants (about two-thirds with a Bachelor's or Master's degree) and contained a gender imbalance (one-third female, two-thirds male), from six European countries, which may restrict the generalizability of the findings (Henrich et al. 2010). Higher education is often associated with stronger numeracy and literacy, which facilitates understanding of structured elicitation tasks (Peters et al. 2006). Gender can also influence preference and susceptibility to heuristics (Croson and Gneezy 2009). Although we found no demographic effects in our ANCOVA analysis, future research should explicitly test whether our findings can be generalized across populations with different educational backgrounds and more balanced gender distributions. Second, the decision problem is simplified, involving only three attributes, which may not fully capture the complexity of real-world decision-making. Moreover, under three attributes, the BWT and OTW tradeoff procedures produce the same weight for the worst attribute, limiting the ability to fully observe the debiasing effects of the BWT method. Future research could address these limitations by investigating decision problems with more than three attributes, both to further explore the debiasing role of BWT and to test whether the anchoring bias persists in more complex and realistic settings. Such extensions would be particularly relevant in high-stakes application areas where biased decisions can lead to substantial negative consequences, such as healthcare, infrastructure planning, or supplier selection. Compared to the traditional Tradeoff procedure, which requires only $N-1$ indifference pairs, BWT requires $2N-3$ for N attributes (Liang et al. 2022). This additional effort, combined with more attributes, may increase cognitive load. It remains an open question whether BWT can still mitigate the anchoring bias effectively under such conditions. Still, BWT remains far less demanding than pairwise comparison approaches, which require $N(N-1)/2$ (unidirectional) or $N(N-1)$ (bidirectional) judgments (Liang et al. 2022). This suggests that BWT may provide a favorable balance between cognitive effort and bias reduction, but its performance in larger scale decision problems should be systematically investigated. In addition, recruiting a more diverse and global participant pool would help assess the robustness of the findings across different populations. Finally, future studies could examine whether similar patterns of the anchoring bias occur across different weight elicitation methods.

Declaration of AI Use

An AI-based language model (ChatGPT 5.2) was used to assist with English language editing.

Author Contributions

Geqie Sun: conceptualization, literature review, experiment design, data collection, methodology, analysis, writing (original draft), editing. **Maarten Kroesen:** conceptualization, methodological guidance, analysis support, editing, supervision. **Jafar Rezaei:** conceptualization, methodological guidance, analysis support, editing, supervision.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study will be openly available in *4TU.ResearchData* after an embargo period.

Endnotes

[1] For simplicity, we use the index j to denote attribute X_j in the remainder of this paper.

[2] For proof of this theorem, please read Wu et al. (2024).

References

Adame, B. J. 2016. "Training in the Mitigation of Anchoring Bias: A Test of the Consider-the-Opposite Strategy." *Learning and Motivation* 53: 36–48.

Beinat, E. 1997. *Value Functions for Environmental Management*. Springer.

Belton, V., and T. Stewart. 2012. *Multiple Criteria Decision Analysis: An Integrated Approach*. Springer.

Buchanan, J. T. 1997. "A Naive Approach for Solving Mcdm Problems: The Guess Method." *Journal of the Operational Research Society* 48, no. 2: 202–206.

Buchanan, J. T., and J. Corner. 1997. "The Effects of Anchoring in Interactive Mcdm Solution Methods." *Computers & Operations Research* 24, no. 10: 907–918.

Bystranowski, P., B. Janik, M. Próchnicki, and P. Skórska. 2021. "Anchoring Effect in Legal Decision-Making: A Meta-Analysis." *Law and Human Behavior* 45, no. 1: 1.

Campbell, S. D., and S. A. Sharpe. 2009. "Anchoring bias in Consensus Forecasts and Its Effect on Market Prices." *Journal of Financial and Quantitative Analysis* 44, no. 2: 369–390.

Chapman, G. B., and E. J. Johnson. 1999. "Anchoring, Activation, and the Construction of Values." *Organizational Behavior and Human Decision Processes* 79, no. 2: 115–153.

Chapman, G. B., and E. J. Johnson. 2002. "Incorporating the Irrelevant: Anchors in Judgments of Belief and Value." *Heuristics and Biases: The Psychology of Intuitive Judgment*: 120–138.

Charness, G., U. Gneezy, and M. A. Kuhn. 2012. "Experimental Methods: Between-Subject and Within-Subject Design." *Journal of Economic Behavior & Organization* 81, no. 1: 1–8.

Choo, E. U., B. Schoner, and W. C. Wedley. 1999. "Interpretation of Criteria Weights in Multicriteria Decision Making." *Computers & Industrial Engineering* 37, no. 3: 527–541.

Croson, R., and U. Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47, no. 2: 448–474.

Edwards, W. 1977. "How to Use Multiattribute Utility Measurement for Social Decisionmaking." *IEEE Transactions on Systems, Man, and Cybernetics* 7, no. 5: 326–340.

Englich, B., T. Mussweiler, and F. Strack. 2006. "Playing Dice With Criminal Sentences: The Influence of Irrelevant Anchors on Experts Judicial Decision Making." *Personality and Social Psychology Bulletin* 32, no. 2: 188–200.

Epley, N., and T. Gilovich. 2005. "When Effortful Thinking Influences Judgmental Anchoring: Differential Effects of Forewarning and Incentives on Self-Generated and Externally Provided Anchors." *Journal of Behavioral Decision Making* 18, no. 3: 199–212.

Fischer, G. W. 1995. "Range Sensitivity of Attribute Weights in Multiattribute Value Models." *Organizational Behavior and Human Decision Processes* 62, no. 3: 252–266.

Fischer, G. W., N. Damodaran, K. B. Laskey, and D. Lincoln. 1987. "Preferences for Proxy Attributes." *Management Science* 33, no. 2: 198–214.

Fishburn, P. C. 1967. "Methods of Estimating Additive Utilities." *Management Science* 13, no. 7: 435–453.

Furnham, A., and H. C. Boo. 2011. "A Literature Review of the Anchoring Effect." *Journal of Socio-Economics* 40, no. 1: 35–42.

Gilovich, T., D. Griffin, and D. Kahneman. 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Uress.

Greenwald, A. G. 1976. "Within-Subjects Designs: To Use or Not to Use?" *Psychological Bulletin* 83, no. 2: 314.

Hämäläinen, R. P., and A. A. Salo. 1997. "The Issue Is Understanding the Weights." *Journal of Multi-Criteria Decision Analysis* 6, no. 6.

Henrich, J., S. J. Heine, and A. Norenzayan. 2010. "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33, no. 2-3: 61–83.

Jaspersen, J. G., and G. Montibeller. 2015. "Probability Elicitation Under Severe Time Pressure: A Rank-Based Method." *Risk Analysis* 35, no. 7: 1317–1335.

Keeney, R. L. 1974. "Multiplicative Utility Functions." *Operations Research* 22, no. 1: 22–34.

Keeney, R. L., and H. Raiffa. 1976. *Decisions With Multiple Objectives: Preferences and Value Trade-Offs*. Cambridge University Press.

Kendall, M. G. 1948. *Rank Correlation Methods*.

Kirkwood, C. W., and R. K. Sarin. 1980. "Preference Conditions for Multiattribute Value Functions." *Operations Research* 28, no. 1: 225–232.

Larichev, O. I. 1992. "Cognitive Validity in Design of Decision-Aiding Techniques." *Journal of Multi-Criteria Decision Analysis* 1, no. 3: 127–138.

Larichev, O. I., and R. V. Brown. 2000. "Numerical and Verbal Decision Analysis: Comparison on Practical Cases." *Journal of Multi-Criteria Decision Analysis* 9, no. 6: 263–273.

Liang, F., M. Brunelli, and J. Rezaei. 2022. "Best-Worst Tradeoff Method." *Information Sciences* 610: 957–976.

Ly, D. P., P. G. Shekelle, and Z. Song. 2023. "Evidence for Anchoring Bias During Physician Decision-Making." *JAMA Internal Medicine* 183, no. 8: 818–823.

McElroy, T., and K. Dowd. 2007. "Susceptibility to Anchoring Effects: How Openness-to-Experience Influences Responses to Anchoring Cues." *Judgment and Decision Making* 2, no. 1: 48–53.

Meub, L., and T. Proeger. 2016. "Can Anchoring Explain Biased Forecasts? Experimental Evidence." *Journal of Behavioral and Experimental Finance* 12: 1–13.

Montibeller, G., and D. von Winterfeldt. 2015. "Cognitive and Motivational Biases in Decision and Risk Analysis." *Risk Analysis* 35, no. 7: 1230–1251.

Montibeller, G., and D. von Winterfeldt. 2024. "Behavioral Decision Research: Descriptive and Prescriptive Perspectives." In *Behavioral Decision Analysis*, 15–40. Springer.

Morewedge, C. K., H. Yoon, I. Scopelliti, C. W. Symborski, J. H. Korris, and K. S. Kassam. 2015. "Debiasing Decisions: Improved Decision Making With a Single Training Intervention." *Policy Insights From the Behavioral and Brain Sciences* 2, no. 1: 129–140.

Morton, A., and B. Fasolo. 2009. "Behavioural Decision Theory for Multi-Criteria Decision Analysis: A Guided Tour." *Journal of the Operational Research Society* 60, no. 2: 268–275.

Mussweiler, T., and F. Strack. 2001. "The Semantics of Anchoring." *Organizational Behavior and Human Decision Processes* 86, no. 2: 234–255.

Mussweiler, T., F. Strack, and T. Pfeiffer. 2000. "Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility." *Personality and Social Psychology Bulletin* 26, no. 9: 1142–1150.

Nagtegaal, R., L. Tummers, M. Noordegraaf, and V. Bekkers. 2020. "Designing to Debias: Measuring and Reducing Public Managers Anchoring bias." *Public Administration Review* 80, no. 4: 565–576.

Palan, S., and C. Schitter. 2018. "Prolific. ACA Subject Pool for Online Experiments." *Journal of Behavioral and Experimental Finance* 17: 22–27.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1992. "Behavioral Decision Research: A Constructive Processing Perspective." *Annual Review of Psychology*.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press.

Peters, E., D. Västfjäll, P. Slovic, C. Mertz, K. Mazzocco, and S. Dickert. 2006. "Numeracy and Decision Making." *Psychological Science* 17, no. 5: 407–413.

Pöyhönen, M., and R. P. Hämäläinen. 2000. "There Is Hope in Attribute Weighting." *INFOR: Information Systems and Operational Research* 38, no. 3: 272–282.

Rezaei, J. 2015. "Best-Worst Multi-Criteria Decision-Making Method." *Omega* 53: 49–57.

Rezaei, J. 2021. "Anchoring Bias in Eliciting Attribute Weights and Values in Multi-Attribute Decision-Making." *Journal of Decision Systems* 30, no. 1: 72–96.

Rezaei, J. 2022. The Balancing Role of Best and Worst in Best-Worst Method. In *Advances in Best-Worst Method: Proceedings of the Second International Workshop on Best-Worst Method (BWM2021)*, pages 1–15. Springer.

Rezaei, J., A. Arab, and M. Mehregan. 2022. "Equalizing Bias in Eliciting Attribute Weights in Multiattribute Decision-Making: Experimental Research." *Journal of Behavioral Decision Making* 35, no. 2: e2262.

Rezaei, J., A. Arab, and M. Mehregan. 2024. "Analyzing Anchoring Bias in Attribute Weight Elicitation of Smart, Swing, and Best-Worst Method." *International Transactions in Operational Research* 31, no. 2: 918–948.

Saaty, T. L. 1980. "The Analytic Hierarchy Process (AHP)." *Journal of the Operational Research Society* 41, no. 11: 1073–1076.

Singer, E., and C. Ye. 2013. "The Use and Effects of Incentives in Surveys." *ANNALS of the American Academy of Political and Social Science* 645, no. 1: 112–141.

Slovic, P. 1995. "The Construction of Preference." *American Psychologist* 50, no. 5: 364.

Smith, J. E., and J. S. Dyer. 2021. "On (Measurable) Multiattribute Value Functions: An Expository Argument." *Decision Analysis* 18, no. 4: 247–256.

Sun, G., M. Kroesen, and J. Rezaei. 2025. "Anchoring Bias in Value Function Elicitation Within Multi-Attribute Value Theory." *Decision Analysis* 22, no. 4: 284–304.

Thirsk, L. M., J. T. Panchuk, S. Stahlke, and R. Hagtvedt. 2022. "Cognitive and Implicit Biases in Nurses' Judgment and Decision-Making: A Scoping Review." *International Journal of Nursing Studies* 133: 104284.

Tversky, A., and D. Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157: 1124–1131.

von Nitzsch, R., and M. Weber. 1993. "The Effect of Attribute Ranges on Weights in Multiattribute Utility Measurements." *Management Science* 39, no. 8: 937–943.

von Winterfeldt, D. and W. Edwards. 1993. *Decision Analysis and Behavioral Research*.

Watson, S. R., and D. M. Buede. 1987. *Decision Synthesis: The Principles and Practice of Decision Analysis*. Cambridge University Press.

Weber, M., and K. Borcherding. 1993. "Behavioral Influences on Weight Judgments in Multiattribute Decision Making." *European Journal of Operational Research* 67, no. 1: 1–12.

Weber, M., F. Eisenführ, and D. von Winterfeldt. 1988. "The Effects of Splitting Attributes on Weights in Multiattribute Utility Measurement." *Management Science* 34, no. 4: 431–445.

Wilson, T. D., C. E. Houston, K. M. Etling, and N. Brekke. 1996. "A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents." *Journal of Experimental Psychology: General* 125, no. 4: 387.

Wu, Q., X. Liu, L. Zhou, J. Qin, and J. Rezaei. 2024. "An Analytical Framework for the Best–Worst Method." *Omega* 123: 102974.

Zionts, S., and J. Wallenius. 1983. "An Interactive Multiple Objective Linear Programming Method for a Class of Underlying Nonlinear Utility Functions." *Management Science* 29, no. 5: 519–529.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. Appendix A.docx.