Applied Cognitive Psychology

RESEARCH ARTICLE OPEN ACCESS

All's Eco-Friendly That Ends Eco-Friendly (If Remembered as Such): Memory Processes in Retrospective Judgment of Environmentally Significant Sequences

Patrik Sörqvist[1,2] | Emil Skog[1] | Johanna Heidenreich[1] | John E. Marsh[1,3]

[1]Department of Health, Learning and Technology, Luleå University of Technology, Luleå, Sweden | [2]Department of Building Engineering, Energy Systems and Sustainability Science, University of Gävle, Gävle, Sweden | [3]Human Factors Laboratory, School of Psychology and Computer Sciences, University of Central Lancashire, Preston, UK

Correspondence: Patrik Sörqvist (patrik.sorqvist@hig.se)

Received: 19 August 2024 | **Revised:** 14 April 2025 | **Accepted:** 17 July 2025

Funding: This work was supported by Riksbankens Jubileumsfond, P23-0067.

Keywords: environmental friendliness | primacy effect | recency effect | retrospective judgment | short-term memory

ABSTRACT

Retrospective judgments of environmentally significant sequences are biased by recency: sequences ending with an environmentally friendly item are rated as more eco-friendly than otherwise identical sequences with the same item earlier in the list. A corresponding primacy effect is typically absent. This may have applied consequences for how consumers perceive the environmental friendliness of their purchase decisions, for example. The aim of the present investigation was to reach a better understanding of why the recency but not the primacy effect manifests in eco-judgments. We found that the recency effect is just as large when continual distraction takes place between item presentations as when it does not. Moreover, memory for recently presented items was better than that for older items, but a filled retention interval reduced the recency effect in both memory and retrospective judgments. These findings support a memory-based explanation of the recency effect in retrospective judgments and suggest that poor memory of items early in the sequence is the reason why the primacy effect in judgments does not manifest.

1 | Introduction

When consumers shop—whether in stores or online—they typically make choices in sequence, selecting one item after another. Consequently, their retrospective evaluation of the purchased items' environmental friendliness is disproportionately shaped by the final item in the sequence, even when the entire episode unfolds over just a few seconds and judgments are immediate (Sörqvist et al. 2024b). Such recency-based distortions in environmental evaluation have important implications. For example, consumers who perceive their shopping behaviour as relatively eco-friendly may subsequently feel licensed to choose less sustainable options, such as carbon-intensive products (Sörqvist et al. 2024a)—an example of negative behavioural spillover (Gholamzadehmir et al. 2019). In this way, memory-based biases may play a direct role in climate-relevant decisions, thereby highlighting a broader role for memory and learning in sustainability efforts (cf. Zhao et al. 2024). The present study investigates why recency effects arise in retrospective judgments of environmental impact, and why a corresponding primacy effect—greater influence from the first item—is typically absent. We test whether this asymmetry reflects differences in how well earlier versus later items are remembered. Our overarching aim is to clarify how memory processes influence perceptions of past behaviour and, by extension, may influence future environmentally relevant decisions.

Recency effects, whereby recently presented items exert a disproportionate influence on evaluation, have been documented across many domains (e.g., vacations, childbirth, noise exposure; see Alaybek et al. 2022; Aldrovandi et al. 2015; Asutay et al. 2021; Chajut et al. 2014; Fredrickson 2000; Fredrickson and Kahneman 1993; Kahneman et al. 1993; Kemp et al. 2008; Montgomery and Unnava 2009; Redelmeier and Kahneman 1996; Robinson et al. 2013; Schreiber and Kahneman 2000). Similar effects have been found in judgments of environmental impact (Sörqvist et al. 2024a, 2024b), while primacy effects—greater influence from the first item—appear more elusive. This may be because recently presented items are more accessible in memory than those encountered earlier.

The absence of a primacy effect is noteworthy given the robust primacy and recency effects observed in short-term memory tasks, where items at the beginning and end of a sequence are typically better remembered than those in the middle. Primacy and recency effects are found in both serial recall (i.e., when items must be recalled in their order of presentation; Cowan et al. 2002; Manning and Schreier 1988) and in free recall (i.e., when participants are free to recall the items in any order; Manning and Schreier 1988; Tan et al. 2010; Ward 2002). Moreover, primacy effects are often stronger than recency effects (Penny 1989). Memory effects in retrospective eco-judgments thus appear to be functionally distinct from the classic short-term memory effects in word-list recall.

Although recency/ending effects are more prevalent than primacy effects in the context of retrospective judgments (Aldrovandi et al. 2015; Fredrickson 2000; Sörqvist et al. 2024b), there are still instances where primacy effects appear more robust (Collins and Shanks 2002). For instance, when asked to pick the most preferred wine from a tasted sequence, people tend to choose the first one they tasted (Mantonakis et al. 2009); information about cars presented first has a larger effect on preference estimates than information presented last (Rey et al. 2019); primacy effects can sometimes be just as large in magnitude as recency effects in the subjective evaluation of visited destinations during a vacation (Zare and Pearce 2022; but see Zare and Pearce 2018); and primacy effects can be stronger than recency effects on product evaluations when interruptions are introduced during the presentation of the to-be-evaluated material (Xu et al. 2023). This motivated an investigation into the underpinning memory processes to better understand why the recency effect but not the primacy effect manifests in retrospective judgments.

To clarify the memory mechanism behind the recency effect, we introduced a continual distraction task; a task sandwiched in-between each to-be-recalled item and following the last item during sequence presentation (e.g., Bjork and Whitten 1974; Tzeng 1973). Continual distraction usually increases the magnitude of the recency effect in word list recall, an effect referred to as the long-term recency effect (Glenberg et al. 1983; Howard et al. 2005; Howard and Kahana 1999; Thapar and Greene 1993; Watkins et al. 1989). However, when the memory task involves recognition rather than recall, the recency effect seems immune to continual distraction (Talmi and Goshen-Gottstein 2006). Whether retrospective judgments are influenced by continual distraction during item presentation is an open question.

Experiment 1 set out to test the effect of continual distraction on the recency effect in retrospective eco-judgments. One possibility is that the recency effect (i.e., a higher environmental-friendliness estimate for sequences ending with a green-labeled item compared to identical sequences wherein the green-labeled item appeared first; and conversely, a lower environmental-friendliness estimate for sequences ending with a red-labeled item in comparison with other sequences that are identical content-wise but with the red-labeled item at the beginning instead) manifests because items encountered after the first item overwrite or interfere with memory of the first item (cf. Oberauer et al. 2012). Another possibility is that memory of the first item is poorer due to decay (Brown 1958; Peterson and Peterson 1959), again resulting in a recency effect. If so, the recency effect should be larger in trials with continual distraction during item presentation, because decay and interference should foremost impair memory of the (oldest) items presented first in the sequence. Alternatively, the recency effect may reflect temporal distinctiveness (Brown et al. 2007). Such accounts appeal to the general principle of the ratio rule (e.g., Bjork and Whitten 1974) whereby the magnitude of the recency effect depends on the temporal spacing between items. If so, continual distraction should have no effect on memory, and therefore no effect on judgments.

Continual distraction after the final to-be-estimated item has been presented also has the consequence of delaying the judgment. This introduces a short retention interval, which may also influence the recency effect. A filled retention interval between list presentation and test tends to reduce the magnitude of the recency effect for recognition (Talmi and Goshen-Gottstein 2006) as for recall (Bjork and Whitten 1974; Glanzer and Cunitz 1966; Postman and Phillips 1965; Tan and Ward 2000). The same tendency is found in the context of retrospective judgments: a filled retention interval between word presentation and retrospective pleasantness ratings leads to a smaller recency effect in ratings as well as word memory (Aldrovandi et al. 2015). Because of this, we might expect a reduced recency effect with continual distraction that follows after the final item. The role of retention intervals is explored further in Experiment 2.

2 | Experiment 1: Continual Distraction

2.1 | Method

2.1.1 | Participants

We based our determination of the required sample size on an a priori power analysis (using G*Power; Faul et al. 2007) of the effect size of the recency effect (i.e., the difference in eco-judgments of sequence beginning with an environmentally significant item and sequences ending with an environmentally significant item) found in Experiment 1 (with a sample of 27 participants) of the paper by Sörqvist et al. (2024b). A power analysis using Cohen's $d_z = 0.64$ (or $\eta_p^2 = 0.30$), based on Sörqvist et al. (2024b), indicated that a sample of 28 participants would suffice (one-tailed). The recency effect was found with a sample of 32 participants in Experiment 1 of Sörqvist et al. (2024a). Therefore, it was determined that a sample of 32 participants would be appropriate in the current study. Hence, a total of 32 participants (20 female and 12 male, Mean age = 28.6 years, SD = 7.10, comprising students from various European nationalities) took part in

Experiment 1. One participant was replaced due to a technical error. All participants had normal or corrected-to-normal vision and reported no red-green colour blindness. Participants were recruited at a university campus and they received a small honourarium for their participation. The third author recruited participants by asking people face-to-face in the corridor whether they wished to participate, or by posting an announcement on the university's digital platform. Written informed consent was obtained from all participants. This research project received research ethical clearance from the Swedish Ethical Review Authority (Dnr 2023–01109-01). None of the experiments were preregistered.

2.1.2 | Materials

The experiment was created with PsychoPy (Peirce et al. 2019). The materials were modelled after previous research (Sörqvist et al. 2024a). Pictures with pairs of food items (fruit, meat, packages of rice, pasta, etc.) were used as stimulus material (see Supporting Information for examples). One food item was shown on the left and the other item on the right side of the picture. Each food item was given one of the following three eco-labels: a green label with the text message "eco impact A," a yellow label with the text message "eco impact D" or a red label with the text message "eco impact G". The labels were displayed adjacent to the food item. A total of 42 food items were used to create 21 food-pair pictures. Each picture included food items with the same label (i.e., either two items with a green label, two items with a yellow label, or two items with a red label). Thirty food items were labeled with "eco impact D" and used to create 15 pictures with yellow-labeled pairs. The other 12 food items were used to create 6 pictures with items with green labels and 6 pictures with the exact same items but with red labels instead. Thus, the food identity was matched between the two types of labels. This was done to be able to isolate the effect of the label/carbon footprint independently of the properties of the food item. The size of the green and the red eco-label was 1.5 times the size of the yellow eco-label, to make them visually distinct from the yellow labels by both colour and size, which could potentially increase the magnitude of the recency effect (cf. Bornstein et al. 1995).

These stimuli were then used to create five types of shopping sequences: (1) sequences containing only pictures with yellow-labelled food items ("control"), (2) sequences containing one picture with green-labelled food items followed by pictures with yellow-labelled food items ("green primacy"), (3) sequences containing pictures with yellow-labelled food items followed by one picture with green-labelled food items ("green recency"), (4) sequences containing one picture with red-labelled food items followed by pictures with yellow-labelled food items ("red primacy"), and (5) sequences containing pictures with yellow-labelled food items followed by one picture with red-labelled food items ("red recency"). Furthermore, the shopping sequences comprised either three, five, or seven pictures. There were thus a total of 15 different sequences (i.e., 5 sequence types $\times$ 3 lengths).

The trials in the experiment were divided into two blocks. In one of the blocks, participants were presented with a distraction task during sequence presentation. For this distraction task, 75 mathematical statements like "Is $14+25>38$?" and "Is $25+24<49$?" were created. The order of the stimuli, sequence type, and length of the shopping sequences were randomly presented for each participant.

2.1.3 | Design and Procedure

Participants sat in front of a laptop computer in a designated laboratory during the whole data collection. Four participants could take part at the same time. Before beginning the experimental task, each participant read the informed consent form and agreed to participate. After that, they answered a few demographical questions before detailed instructions followed. The data collector (Author three) was available to answer any questions that arose as participants read the instructions. The participants were told that they would be conducting fictitious shopping sequences in which they had to choose between eco-labeled goods (a task modelled after previous research; Sörqvist et al. 2024a). The purpose of this was to simulate decisions consumers make during a shopping sequence while maintaining experimental control of the presentation of products and labels. The eco-labelling was introduced as being based on a life cycle analysis where the goods' environmental impact was calculated in the entire chain from production to recycling including transportation. They were also told that after each shopping sequence they would be requested to estimate the environmental friendliness of the whole shopping sequence and they were told about the number of trials in total.

Between the two trial blocks, participants got the opportunity to take a self-paced pause. Each block comprised 3 instances of each sequence length and sequence type, resulting in 45 shopping sequences/trials in each block. In addition to these 45 shopping sequences, 3 warmup trials were conducted at the beginning of each block. Responses from the warmup trials were excluded from the data analysis. In one of the two trial blocks (the distraction block), participants conducted three tasks: selecting the food item they would prefer to buy from each pair, conducting a mathematical distraction task after each food selection (i.e., answering "yes" or "no" to a mathematical statement such as "Is $14+25>38$?"), and estimating how environmentally friendly the most recent shopping sequence was. In the other trial block (the no-distraction block), the participants conducted two tasks: selecting the food item they would prefer to buy from each pair, and estimating how environmentally friendly the most recent shopping sequence was. The order of the two blocks were counterbalanced between participants.

Each shopping sequence began with the words "shopping sequence X" displayed on the screen. "X" was replaced by the respective trial number. The participant started the shopping sequence by pressing the space bar. Immediately after the button press, a picture with a pair of food items was displayed and the participants were asked to choose one of the two by pressing either the left or the right arrow key on the keyboard. When they were in the distraction block, they then saw a mathematical statement (e.g., "Is $12+16>28$?") and asked to press the left arrow key to answer "no" to the statement or the right arrow key to answer "yes" to the statement. If they were not in the distraction block, the computer

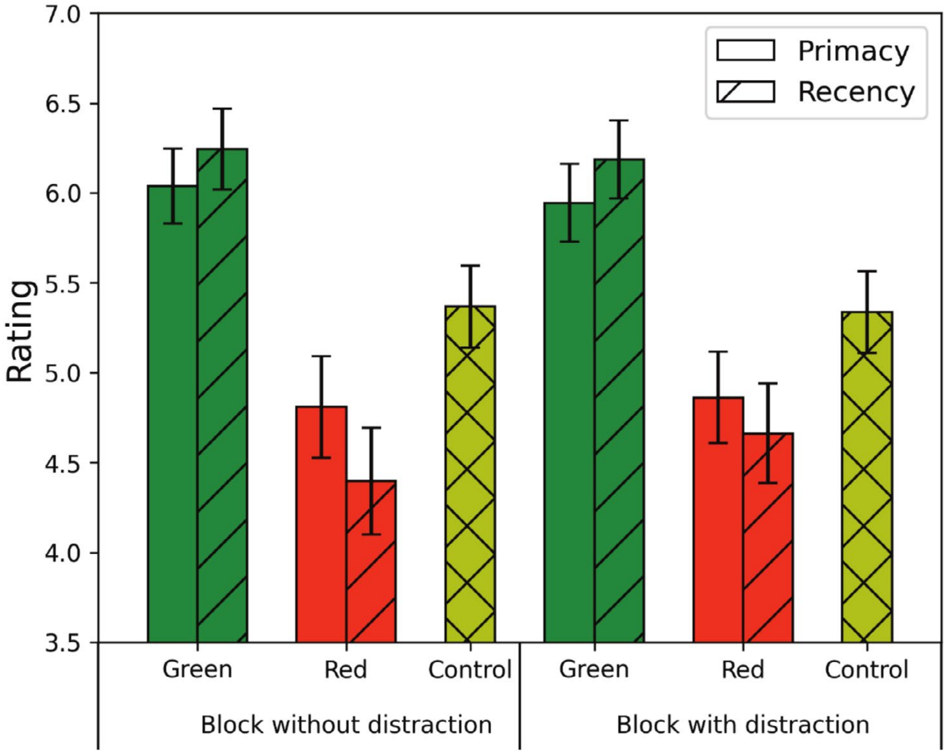

FIGURE 1 | Mean estimates of the environmental friendliness of shopping sequences in Experiment 1 during trials blocks with versus without continual distraction. The sequences comprised a red (environmentally harmful) item first or last or a green (environmentally friendly) item first or last or had yellow (intermediate environmental footprint) items only (control). Error bars represent standard error of means.

immediately presented another food-pair picture instead. This procedure was repeated until the full length of the sequence was reached. After the last good selection (or the last distraction task, when participants were in the distraction block) participants were presented with the statement "This shopping sequence was environmentally friendly" and they were asked to indicate their level of agreement with the statement on a scale from 1 (not at all) to 9 (completely) by pressing the corresponding number key. Then the next shopping sequence started. The whole experiment took about 40 min to complete.

In sum, the experiment comprised a 5 (type of shopping sequence: control, green primacy, green recency, red primacy, red recency) × 3 (shopping sequence length [number of good selections]: three, five, seven) × 2 (block: with or without distraction task) factorial within-subjects design. The rating of environmental-friendliness was used as dependent variable.

2.1.4 | Analysis

We set our alpha to 0.05 in all experiments below. The statistical analyses were performed using IBM SPSS version 27 and JASP version 0.17.3.0. Bayesian factors (with JASP's default Bayesian priors) are reported for all statistical tests and expressed as BF_{10} values, which express how many times more likely the hypothesis is over the null-hypothesis. Lee and Wagenmakers' (2014) conventional interpretation scheme categorizes Bayes factors between 3 and 10 as moderate evidence for the hypothesis over the null, and factors above 10 as strong evidence. For cases where analyses of variance substantially violate normality and sphericity, Bayes factors diverge substantially from frequentist analyses. In these cases, Bayes factors were not reported as they otherwise would misinform interpretations. We employed this rationale for omitting Bayes factors in certain cases in all four experiments equally.

2.1.5 | Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. No data have been excluded. The data for all experiments reported in this paper can be accessed at https://doi.org/10.17605/OSF.IO/EXHAV (Sörqvist et al. 2024c).

2.2 | Results and Discussion

As can be seen in Figure 1, the results revealed a recency effect for both green (i.e., environmentally friendly) and red (i.e., environmentally harmful) labels across both trial blocks. A sequence ending with the selection of a red item was perceived as less environmentally friendly than sequences that were identical content-wise but wherein the red item was instead selected first in the sequence. Conversely, a sequence ending with the selection of a green item was perceived as more environmentally friendly than identical sequences with the green item first. These recency effects were similar in magnitude when participants performed a distraction task (sandwiched between item selections during sequence presentation) and when there was no distraction task. The conclusions were reinforced by a repeated measures 2 (block: with vs. without distraction) × 2 (colour: with green item vs. with red item) × 2 (sequence type: primacy [red/green item first] vs. recency [red/green item last in the sequence]) analysis of variance with estimates of environmental friendliness as the dependent variable. The analysis revealed a significant effect of colour, $F = 42.36$, $p < 0.001$, $\eta_p^2 = 0.58$, $BF_{10} = 270604.41$, a significant interaction between colour and sequence type, $F = 16.30$, $p < 0.001$, $\eta_p^2 = 0.35$, $BF_{10} = 36.17$, but no other significant interactions or main effects. Most importantly, the three-way interaction was nonsignificant, $F = 0.56$, $p = 0.460$, $\eta_p^2 = 0.02$, $BF_{10} = 0.24$ (indicating support for the null-hypothesis). The effect size of the recency effect was largest for red colours in the block without distraction (Cohen's $d_z = 0.64$, $t[31] = 3.60$, $p < 0.001$, one tailed, $\eta_p^2 = 0.30$), smallest for green colours in the block without distraction (Cohen's $d_z = 0.30$, $t[31] = 1.70$, $p = 0.05$, one tailed, $\eta_p^2 = 0.09$), and intermediate for green colours in the block with distraction (Cohen's $d_z = 0.41$, $t[31] = 2.34$, $p = 0.013$, one tailed $\eta_p^2 = 0.15$) and for red colours in the block with distraction (Cohen's $d_z = 0.34$, $t[31] = 1.90$, $p = 0.034$, one tailed, $\eta_p^2 = 0.11$). Given our a priori estimate of required sample size and the current results, we conclude that the current experiment was appropriately powered with $N = 32$. For control purposes, an analysis of time spent on the distraction task was conducted. The average time spent was 5.41 s/mathematical problem (SD = 2.04). Three participants seemed to have skipped through the distraction task part of the experiment, because their response times were outstandingly short (less than 3 s/problem [0.79–2.56 s/problem]). Control analyses with these participants removed did not change the conclusions from the experiment. The three-way interaction still showed no effect of the distraction task, $F = 0.27$, $p = 0.607$, $\eta_p^2 = 0.01$, $BF_{10} = 0.27$. Control analyses also revealed that sequence length (3 vs. 5 vs. 7 item selections) did not interact with the other independent variables. Thus, the distraction task had no effect, regardless of sequence length.

Taken together, Experiment 1 replicated the recency effect in retrospective eco-estimates of item sequences seen in previous research (Sörqvist et al. 2024a, 2024b) and it was comparably small in magnitude despite the more visually distinct green/red labels. The novel finding reported here is that the effect appears to withstand influence from continual distraction. If decay or interference is what produces the recency effect, by operating primarily on memory of the first list items, then the recency effect should have been larger with continual distraction, since continual distraction should lead to greater loss of the memory record of the initial list item(s). Since this was not the case, the results are more in line with a temporal distinctiveness account of the recency effect, as the relative accessibility of the first and the last sequence item remains the same, regardless of the presence or absence of continual distraction.

3 | Experiment 2: Delayed Judgment After a Filled Retention Interval

The recency effect in eco-judgments survived the short delay of judgment introduced by continual distraction after the final to-be-estimated item. However, previous research shows that a longer filled retention interval between word presentation and retrospective pleasantness ratings leads to a smaller recency effect in ratings as well as word memory (Aldrovandi et al. 2015). If the recency effect in eco-judgments is found because people have better memory of recently presented items, then a longer filled retention interval should reduce the recency effect in judgments, because it puts the most recently presented item at a relatively large disadvantage. Also, estimates of sequences that include an environmentally significant item (e.g., a sequence that *ends* with a high carbon footprint item but otherwise comprises items with an intermediate carbon footprint only) should converge with those for sequences that include no environmentally significant items (e.g., a sequence with only intermediate carbon footprint items), because the lost memory record of the items should make participants resort to guessing. Experiment 2 aimed to test these hypotheses.

Experiment 2 also offered an opportunity to test the assumption that temporal distinctiveness is the mechanism responsible for the recency effect. If temporal distinctiveness rather than interference or decay is responsible for the recency effect, then the filled retention interval should reduce the recency effect because, with increasing delay, time is psychologically compressed as it passes such that, with the progression of time, items appear closer together, are temporally less distinct, and thus more difficult to retrieve. In turn, if decay or interference is responsible for the recency effect, then the retention interval should increase the recency effect's magnitude (or at least leave it unchanged) because decay and interference should operate largely on early presented items.

3.1 | Method

3.1.1 | Participants

To determine an appropriate sample size for Experiment 2, an a priori power analysis was conducted using the effect size of the effect of a filled retention interval on retrospective judgments reported in Aldrovandi et al. (2015, Experiment 3). This effect size (Cohen's d_z) was 0.49, indicating that a sample of 57 participants should be appropriate for detecting the expected effects and interactions (this effect and sample size gives a $\eta_p^2 = 0.196$). Thus, a total of 60 participants took part in Experiment 2 (45 female and 15 male, Mean age = 34.45 years, SD = 14.34). None of them had taken part in the other experiments of the current series. All participants were English native speakers and were based in the UK; they also reported normal or corrected-to-normal vision and reported no form of color-blindness. Twenty-nine participants were recruited at the University of Central Lancashire, and the remaining thirty-one participants were recruited via Prolific Academic (www.prolific.com; Palan and Schitter 2018). Participants from the University of Central Lancashire were compensated with course credits and the participants from Prolific were paid £4.5 for participating in this experiment, which took around 30 min to complete. Informed consent was obtained from all participants.

3.1.2 | Materials, Design, and Procedure

The materials, design and procedure were identical to that of Experiment 1 with the following exceptions. The experiment

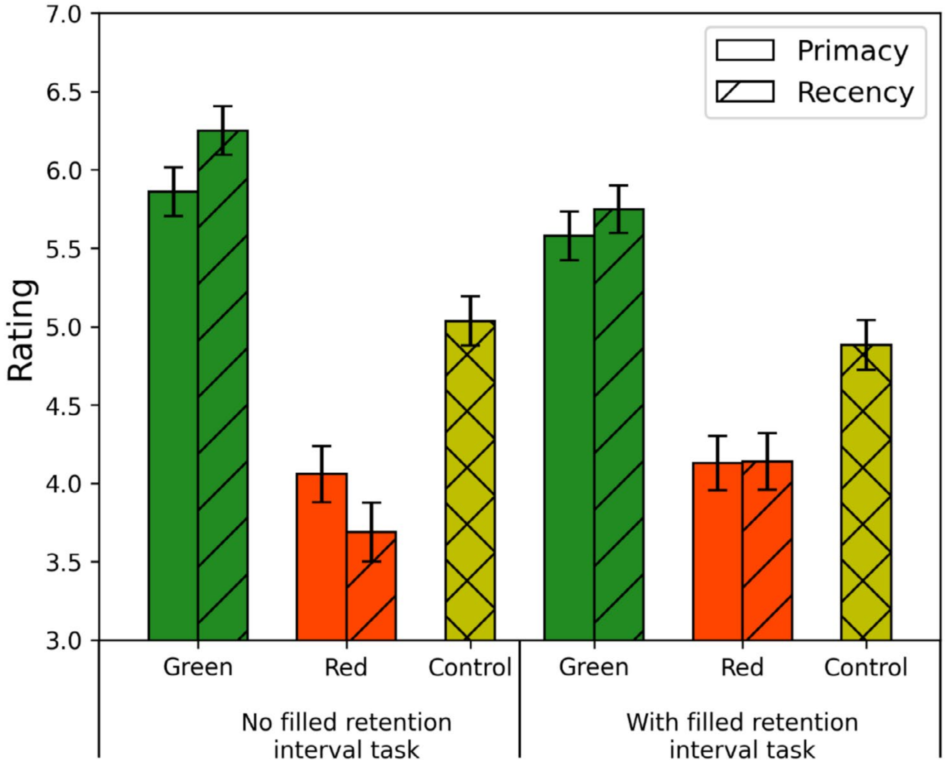

FIGURE 2 | Mean estimates of the environmental friendliness of shopping sequences in Experiment 2 for trials with versus without a filled retention interval. The sequences comprised a red (environmentally harmful) item first or last or a green (environmentally friendly) item first or last or had yellow (intermediate environmental footprint) items only (control). Error bars represent standard error of means.

was created with PsychoPy (PsychoJS, version 2023.2.3; Peirce et al. 2019) and JavaScript to run on the online experiment delivery platform Pavlovia (www.pavlovia.org). Participants were redirected to the Pavlovia experiment via a hyperlink. Participants were only allowed to participate on desktop computers, and they used their own device in an uncontrolled environment.

The experiment comprised two blocks with 30 trials in each block, for a total of 60 trials. The two blocks were identical and separated by a self-paced pause. Half of the trials in each block (15 trials) had a retention interval between list presentation and judgment. The other half (15 trials) were similar but without the retention interval. The order of trials with versus without retention interval was random. In trials with a retention interval, participants were presented with a mental arithmetic task during the retention interval. The task was to calculate the sum of numbers presented on the computer screen. A first two-digit number (randomly selected from the set 10–99) was presented 0.25 s after the final item at list presentation. This two-digit number was followed by another two single-digit numbers, presented for 3 s each, with a blank interstimulus interval of 0.25 s. When the final digit disappeared, the participants were requested to report the sum of the numbers by typing on the keyboard. When done, they pressed a button and made the judgment of the most recent shopping sequence's environmental friendliness, the same way as in the block without a retention interval. In sum, the experiment comprised a 5 (type of shopping sequence: control, green primacy, green recency, red primacy, red recency) × 3 (shopping sequence length: three, five, seven) × 2 (trial type: with or without retention interval) factorial within-subjects design. As in Experiment 1, the rating of environmental-friendliness was used as the dependent variable.

3.2 | Results and Discussion

As can be seen in Figure 2, the results showed that the expected recency effect was present for both green (i.e., environmentally friendly) and red (i.e., environmentally harmful) labels in the trials without a filled retention interval task (left in Figure 2). However, the recency effect disappeared or was diminished in the trials wherein the participants carried out the filled retention interval task (right in Figure 2). This key difference was verified by a repeated measures 2 (task: with vs. without task) × 2 (colour: green vs. red items) × 2 (sequence type: primacy [red/green items first] vs. recency [red/green items last in the sequence]) × 3 (sequence length: 3 vs. 5 vs. 7 items in the list) analysis of variance with estimates of environmental friendliness as the dependent variable. To limit the scope of this four-factor analysis, we first looked at the potential effect of sequence length, expecting that this would have no effect. There was indeed no main effect of sequence length, $F = 1.85$, $p = 0.162$, $\eta_p^2 = 0.030$, $BF_{10} = 0.032$, nor interactions with other factors, except for colour, $F = 3.56$, $p = 0.032$, $\eta_p^2 = 0.057$, $BF_{10} = 0.137$, which we do not interpret as theoretically meaningful, especially since the Bayes factor favours the null-hypothesis. For the other three more important factors, the analysis revealed a significant effect of colour, $F = 134.92$, $p < 0.001$, $\eta_p^2 = 0.696$, $BF_{10} = \infty$, a significant interaction between colour and sequence type, $F = 26.59$, $p < 0.001$, $\eta_p^2 = 0.311$, $BF_{10} = 3.637 \times 10^6$, a significant interaction between task and colour, $F = 35.84$, $p < 0.001$, $\eta_p^2 = 0.378$, $BF_{10} = 9.961 \times 10^7$, and, most importantly, a significant three-way interaction between task, colour, and sequence type, $F = 16.19$, $p < 0.001$, $\eta_p^2 = 0.215$, $BF_{10} = 21986.989$. There were no other significant main effects or interactions. We conclude that our choice of sample size ($N = 60$), based on the effect size reported in Aldrovandi et al. (2015, Experiment 3), was appropriate for detecting the above three-way interaction.

The three-way interaction between task, colour, and sequence type was not significant in Experiment 1 with the continual distraction task but was significant in Experiment 2 with the filled retention interval task. This confirms the hypothesis that the recency effect is diminished when judgments of environmental friendliness follow a filled retention interval task, compared to when judgments follow immediately after sequence presentation (Aldrovandi et al. 2015). Figure 2 further shows that ratings for trials with the filled retention interval task are closer to the mean than ratings in trials without that task. This suggests an effect on memory whereby the filled retention interval task makes participants less likely to remember a red or green item, regardless of where the item appears in the sequence. Thus, the hypothesis that the presence of a task-filled retention interval should reduce the magnitude of the recency effect was confirmed. Supported also was the second hypothesis that the task-filled retention interval makes estimates of sequences that include an environmentally significant item more comparable to that of control sequences.

4 | Experiment 3: Memory of To-Be-Estimated Items

If retrospective judgments are underpinned by the accessibility in memory of items at the point of decision, then a reduced recency effect in judgments should be accompanied by poorer memory for recently presented items. Experiments 3–4 explore this link more directly by asking participants to recall the coloured labels associated with probed items in different positions in the sequences. In Experiment 3 (the first study wherein memory for the to-be-estimated items was measured directly), the participants were asked to make the estimates immediately following item presentation (similar to trials in Experiment 2 without a filled retention interval) and thereafter their memory of the to-be-estimated items was tested.

4.1 | Method

4.1.1 | Participants

As an approach to determine an appropriate sample size for Experiment 3, we conducted an a priori power analysis using the effect sizes of the recency effects for recall reported in Talmi and Goshen-Gottstein (2006) as their probed recall paradigm had a resemblance to the probed recall paradigm used in Experiment 3. The recency effects in their experiment had an effect size of Cohen's d_z of 0.78 and 0.49 ($\eta_p^2 = 0.382$ and $\eta_p^2 = 0.196$), respectively (as calculated from the reported t-values and sample size, using a formula in Lakens 2013). The more conservative of the two suggested that a sample size of 57 participants would be appropriate to detect the expected effects and interaction. Consequently, a group of 60 participants (37 female and 23 male, Mean age = 41.45 years, SD = 12.61) was recruited for Experiment 3. This group will henceforth be called the source memory group. Another 60 participants (35 female and 25 male, Mean age = 45.22 years, SD = 14.99) were recruited for a control group. This served the purpose of controlling for the potential effect of the source memory task (see description below) on eco-ratings. All participants were English native speakers and were based in the UK; they also reported normal or corrected-to-normal vision and reported no form of colour-blindness. All participants were recruited via Prolific and were paid £4.5 for participating. Informed consent was obtained from all participants.

4.1.2 | Materials

The materials were identical to that of Experiment 2 unless noted. The task was simplified, so that only one picture was presented at a time. Thus, participants no longer made active choices between products at item presentation, but instead passively viewed products at presentation. This change was done to increase experimental control of stimulus presentation, task pace and the involved cognitive processes. A total of 57 pictures comprised items with yellow (eco impact D) labels. Another 27 pictures were created with items with green (eco impact A) labels. A further set of 27 pictures was created with the same items as the ones with green labels but with red (eco impact G) labels instead. Thus, the food identity was matched between the green and red types of labels. These stimuli were then used to create the same type of sequences as in Experiment 2 ("control," "green primacy," "green recency," "red primacy," and "red recency"). In addition, "green/red middle sequences" were created, with the green/red item presented in the middle. All sequences had seven items each.

4.1.3 | Design and Procedure

Before beginning the experiment, participants read the informed consent form and agreed to participate. Detailed instructions followed. The participants were told that they would be viewing sequences of food products with various eco-labels. The eco labels, trial structure, and the tasks were also explained to the participants. Participants were encouraged to try to respond quickly and accurately.

When participants had read the instructions, they pressed a button to proceed to the task. Each trial had three phases (Figure 3): a presentation phase, a judgment phase, and a source memory phase (for the source memory group) or a mood-rating phase (for the control group). The participants pressed a button to initiate the presentation phase. At this phase, a sequence of seven items was presented at the center of the computer screen. Each item was displayed for 1500 ms and the items were separated by a 500 ms blank interstimulus interval. The computer randomly selected the type of sequence to be presented and filled the sequences with items by randomly selecting pictures from

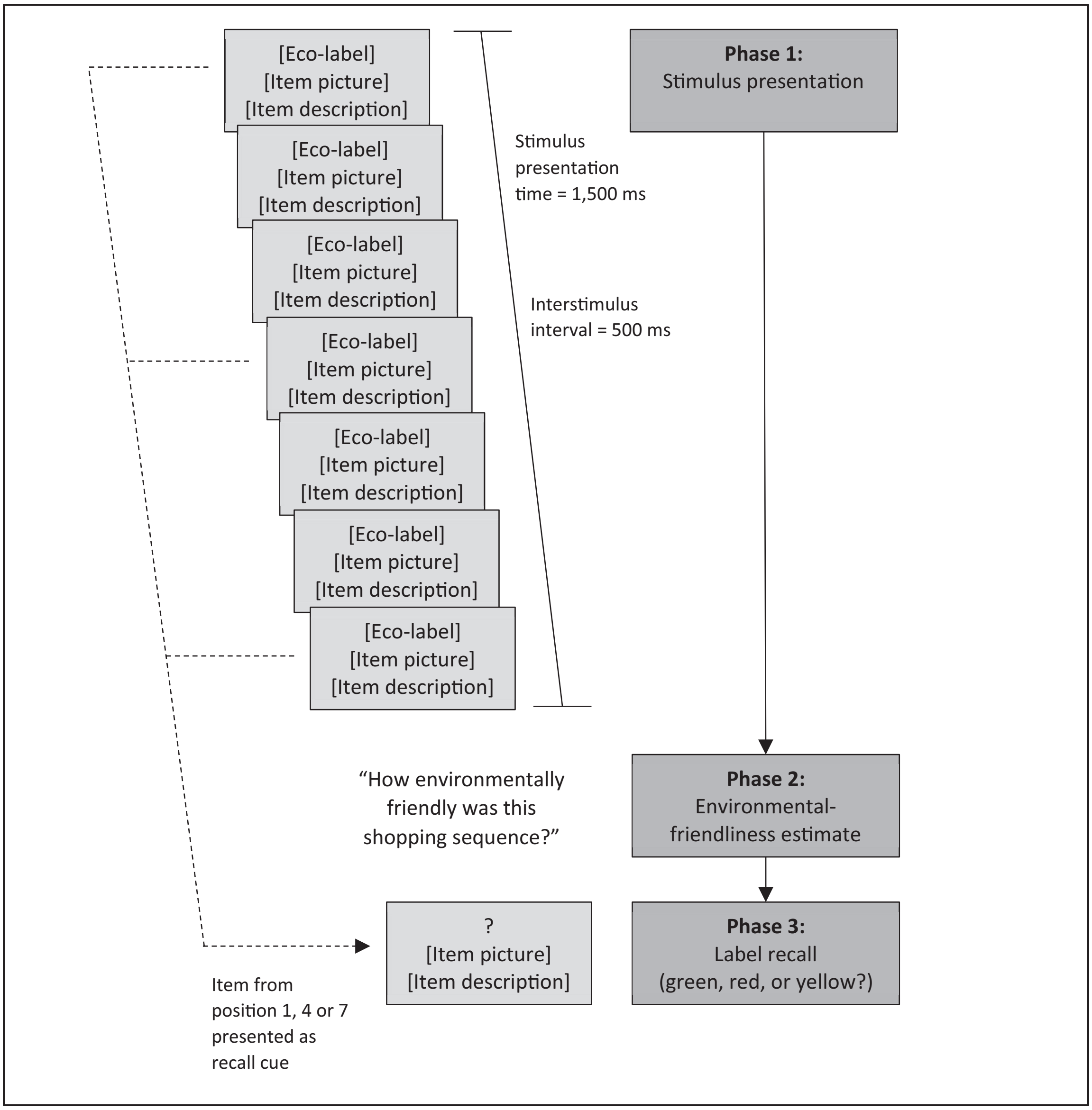

FIGURE 3 | The figure shows the trial structure of Experiment 3.

the picture pool, with the restriction that each type of sequence should be presented 9 times each across the full experiment.

Immediately after the final item had been displayed, the trial moved into the next phase wherein participants were asked to estimate the items' environmental friendliness. Estimates were made on a scale from 1 (not at all environmentally friendly) to 9 (very environmentally friendly) by pressing the corresponding number key on the keyboard.

As participants pressed a number key to make their estimate, the trial moved on to the final phase. For the source memory group, one of the items from the last sequence was presented at the center of the computer screen (the food item and its text description) but without the label. The participants were also presented with three response options; buttons "'r," "g," and "y," could be pressed to identify if the item's label was red, green, or yellow, respectively. The program selected the presented item from list position 1 (first), 4 (middle) or 7 (last). In the control group, the source memory task ("phase 3" in Figure 3) was replaced by a task wherein the participants rated their current mood, a task that was unrelated to the to-be-estimated items and imposed less (or no) demand on memory. Participants rated their own current mood on a three-point scale where 1 = Bad mood, 2 = Neutral, and 3 = Good mood, using the 1–3 number keys on their keyboard. The mood-rating task was considered a filler task, and its results were not analyzed.

The experiment comprised a total of 63 experimental trials (7 sequence types, presented 9 times each, 3 times with the memory cue drawn from list position 1, 3 times with the memory cue drawn from list position 4, and 3 times with the memory cue drawn from list position 7). The experimental trials were preceded by 3 warmup trials that were used to familiarise the participants with the procedure, and the detailed instructions (see above) were repeated before each warmup trial. Data from the warmup trials were not included in the analysis. After having completed 1/3rd, 2/3rds, and 3/3rds of the experiment, participants were shown an "attention check" question which prompted them to click the "z," "x," or "c" keys on their keyboards

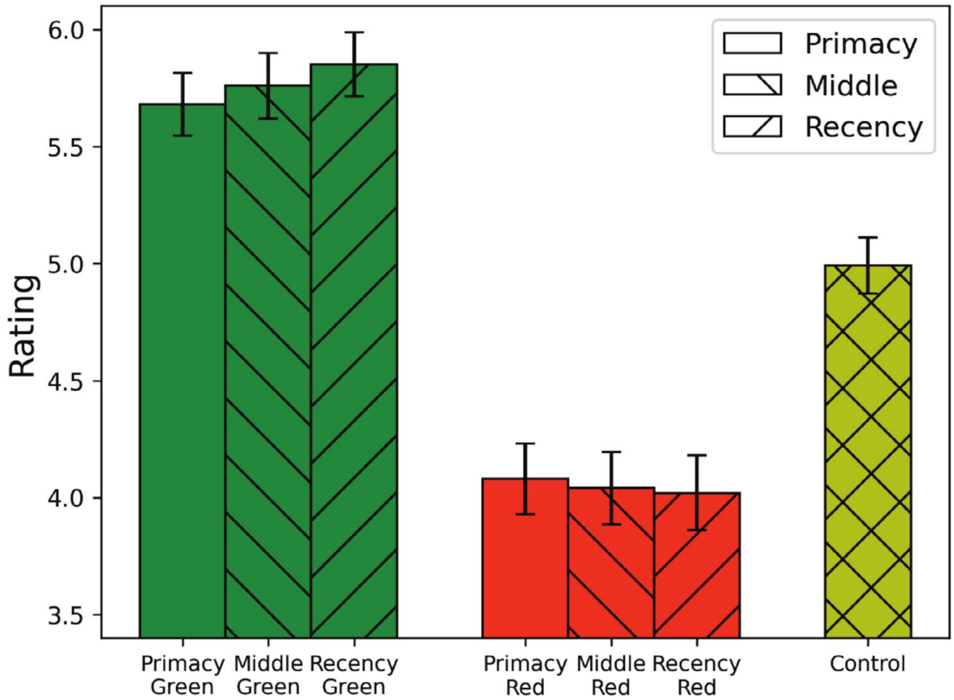

FIGURE 4 | Mean estimates of the environmental friendliness of shopping sequences in the recall group of Experiment 3. The sequences comprised a red (environmentally harmful) item first, in the middle or last, or a green (environmentally friendly) item first, in the middle or last, or they had yellow (intermediate environmental footprint) items only (control). Error bars represent standard error of means.

to continue. Following this, participants were informed of how far they had progressed in the experiment. The experiment took about 30 min to complete.

In sum, the experiment comprised a 7 (type of shopping sequence: control, green primacy, green middle, green recency, red primacy, red middle, red recency) × 3 (cued memory position: 1 vs. 4 vs. 7) factorial within-subjects design. Two dependent measures were used: rating of environmental-friendliness and source memory accuracy. Response time (RT) in the source memory task was also measured and these data are available online (https://doi.org/10.17605/OSF.IO/EXHAV) but are not reported on further.

4.2 | Results and Discussion

Figure 4 shows the results of the rating task in the source memory group and Figure 5 shows the results of the rating task in the control group, respectively. Again, the recency effect was replicated but no evidence of a corresponding primacy effect was found. We first turn our focus to the ratings by the source memory group. A repeated measures 2 (colour: with green item vs. with red item) × 3 (sequence type: primacy [red/green item first] vs. middle [red/green item in the middle] vs. recency [red/green item last in the sequence]) analysis of variance with ratings of environmental friendliness in the source memory group as the dependent variable was conducted. In this analysis, the assumption of sphericity was met (Mauchly's W test), but the rating data for all seven sequence types violated the assumption of normality (Shapiro–Wilk's test). The analysis revealed a significant effect of colour, $F = 129.8$, $p < 0.001$, $\eta_p^2 = 0.687$, $BF_{10} = 1.67 \times 10^{14}$, no significant effect of sequence type, $F = 2.48$, $p = 0.088$, $\eta_p^2 = 0.04$, $BF_{10} = 11.11$ (but the Bayes factor is strongly in favor of the hypothesis over the null), and a significant interaction between colour and sequence type, $F = 6.85$, $p = 0.002$, $\eta_p^2 = 0.104$, $BF_{10} = 56.23$. This outcome is similar to Experiment 1, where product sequences ending with green items lead to more positive ratings of environmental impact compared to product sequences beginning with green items (Figure 4). A similar, but slightly weaker, pattern is seen with red items, but here the ratings are negative, as to be expected (Figure 4). To gain a more direct comparison to Experiment 1, we removed the 'middle' product sequence and revised the above repeated measures analysis of variance to: 2 (colour: with green item vs. with red item) × 2 (sequence type: primacy [red/green item first] vs. recency [red/green item last in the sequence]). Results showed a significant effect of colour, $F = 126.61$, $p < 0.001$, $\eta_p^2 = 0.682$, $BF_{10} = 6.01 \times 10^{15}$, a significant effect of sequence type, $F = 6.08$, $p = 0.017$, $\eta_p^2 = 0.093$, $BF_{10} = 90.59$, and a significant interaction between colour and sequence type, $F = 11.37$, $p = 0.001$, $\eta_p^2 = 0.162$, $BF_{10} = 352.44$. Removing the "middle" product sequence changed the results of the rating task slightly by making the main effect of sequence type significant, and it also increased the effect size of the interaction from $\eta_p^2 = 0.104$ to $\eta_p^2 = 0.162$.

We now turn to the question of whether the source memory task influenced the eco-ratings. Visual inspection of Figures 4 and 5 indicates similarity, suggesting that the source memory task did not influence ratings. If the source memory task had no effect on the judgment task, then there should be no three-way interaction between the two groups, sequence (primacy/middle/recency), and colour (green vs. red) if tested in an analysis of variance. Indeed, this was the case. In this analysis, the assumption of sphericity was violated (Mauchly's W test) for the interaction between colour

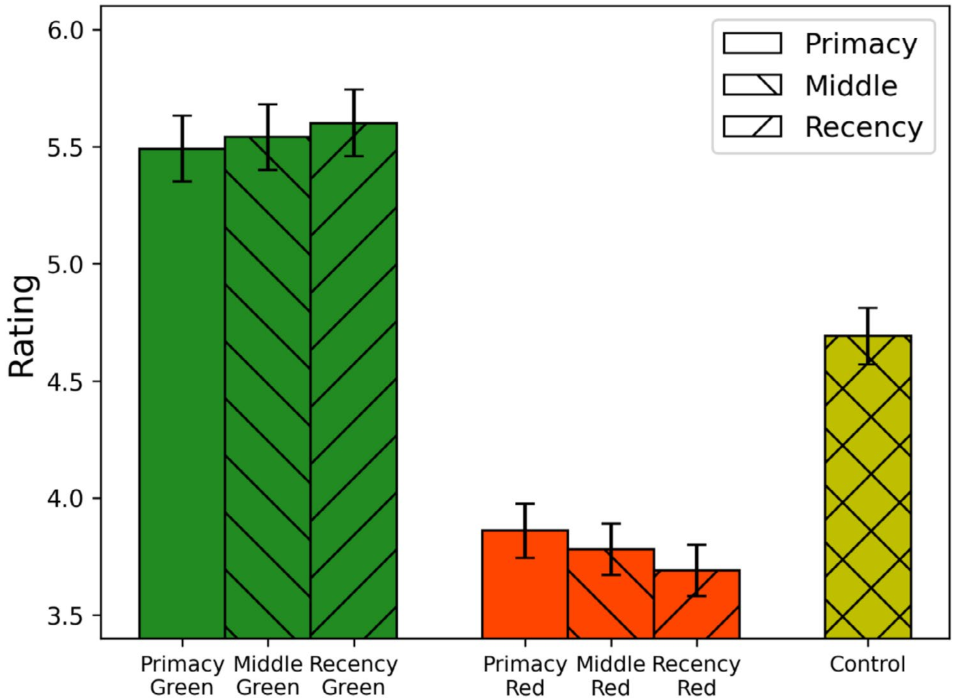

FIGURE 5 | Mean estimates of the environmental friendliness of shopping sequences in the control group of Experiment 3. Data is structured like Figure 4. Error bars represent standard error of means.

and sequence, but Mauchly's W was still high (0.91) and thus was considered acceptable for reporting Bayes factors. The analysis revealed a main effect of colour, $F=338.10$, $p<0.001$, $\eta_p^2=0.741$, $BF_{10}=7.45\times10^{13}$, but no main effect of group, $F=2.24$, $p=0.137$, $\eta_p^2=0.019$, $BF_{10}=0.31$ (indicating evidence in favor of the null-hypothesis), no main effect of sequence, $F=0.23$, $p=0.792$, $\eta_p^2=0.002$, $BF_{10}=31184.33$ (although the Bayes factor was clearly in favor of the hypothesis; see van den Bergh et al. 2023, for a discussion of discrepancies between frequentist and Bayesian repeated-measures analysis). There was a significant interaction between colour and sequence, $F=16.89$, $p<0.001$, $\eta_p^2=0.125$, $BF_{10}=184865.12$, but no significant interaction between colour and group, $F=0.07$, $p=0.792$, $\eta_p^2=0.001$, $BF_{10}=6.35\times10^{-4}$, and no significant interaction between sequence and group, $F=2.82$, $p=0.062$, $\eta_p^2=0.023$, $BF_{10}=0.25$. Most importantly, the three-way interaction was very small and nonsignificant, $F=0.24$, $p=0.788$, $\eta_p^2=0.002$, $BF_{10}=4.95\times10^{-4}$ (indicating strong evidence in favor of the null-hypothesis). We now return to the question of whether there is a relation between memory of the list items and retrospective estimates of those items.

The novelty in Experiment 3 was the use of a source memory task that cued participants to recall the label colour of the first, fourth (middle), or seventh (last) item on each trial, after having completed the rating task. Figure 6 shows the accuracy results in this source memory task. Overall, participants remembered the red and green items with equal accuracy. But, the control (yellow) items were identified with near-perfect accuracy in all three to-be-remembered positions. To interpret these results, it is useful to consider the distinction between source memory and source guessing. Even if participants do not remember the label from presentation, guessing that the label was yellow will result in an accurate response on most trials, since a large majority of items had yellow labels. Because of this, the high performance in the control conditions probably reflects guessing rather than actual source memory.

The most important results in relation to source memory, however, are the differences in memory for green/red labels at different serial positions. Figure 6A shows the full results across the seven types of shopping sequences and the three types of cued recall positions. To aid analysis of recency versus primacy effects in the source memory task, Figure 6B shows more focused results where % correct identification for the red and green sequences has been averaged, and the control sequence has been omitted. The primary comparison of interest is comparing the primacy sequence where the cued recall position was 1 (first item) versus the middle sequence where the cued recall position was 4 (middle item) vs. the recency sequence where the cued recall position was 7 (last item). Focusing on these three cases, Figure 6B shows that the recency sequence where the cued recall position was 7 (last item) produced the highest source memory accuracy, revealing a recency effect but no primacy effect in the source memory task. A repeated measures 3 (cued recall position: 1 vs. 4 vs. 7)×3 (sequence type: primacy [red/green item first] vs. middle [red/green item in the middle] vs. recency [red/green item last in the sequence]) analysis of variance with source memory % accuracy as the dependent variable was used to analyse the results in the source memory task (Figure 6B). The assumption of sphericity was violated (Mauchly's W test), and the data violated the assumption of normality (Shapiro–Wilk's test). The analysis was thus corrected with a Huynh–Feldt correction, and no Bayes factors are reported. Results revealed no significant effect of cued recall position, $F=1.91$, $p=0.156$, $\eta_p^2=0.031$, no significant effect of sequence type, $F=1.02$, $p=0.360$, $\eta_p^2=0.017$, but a significant interaction

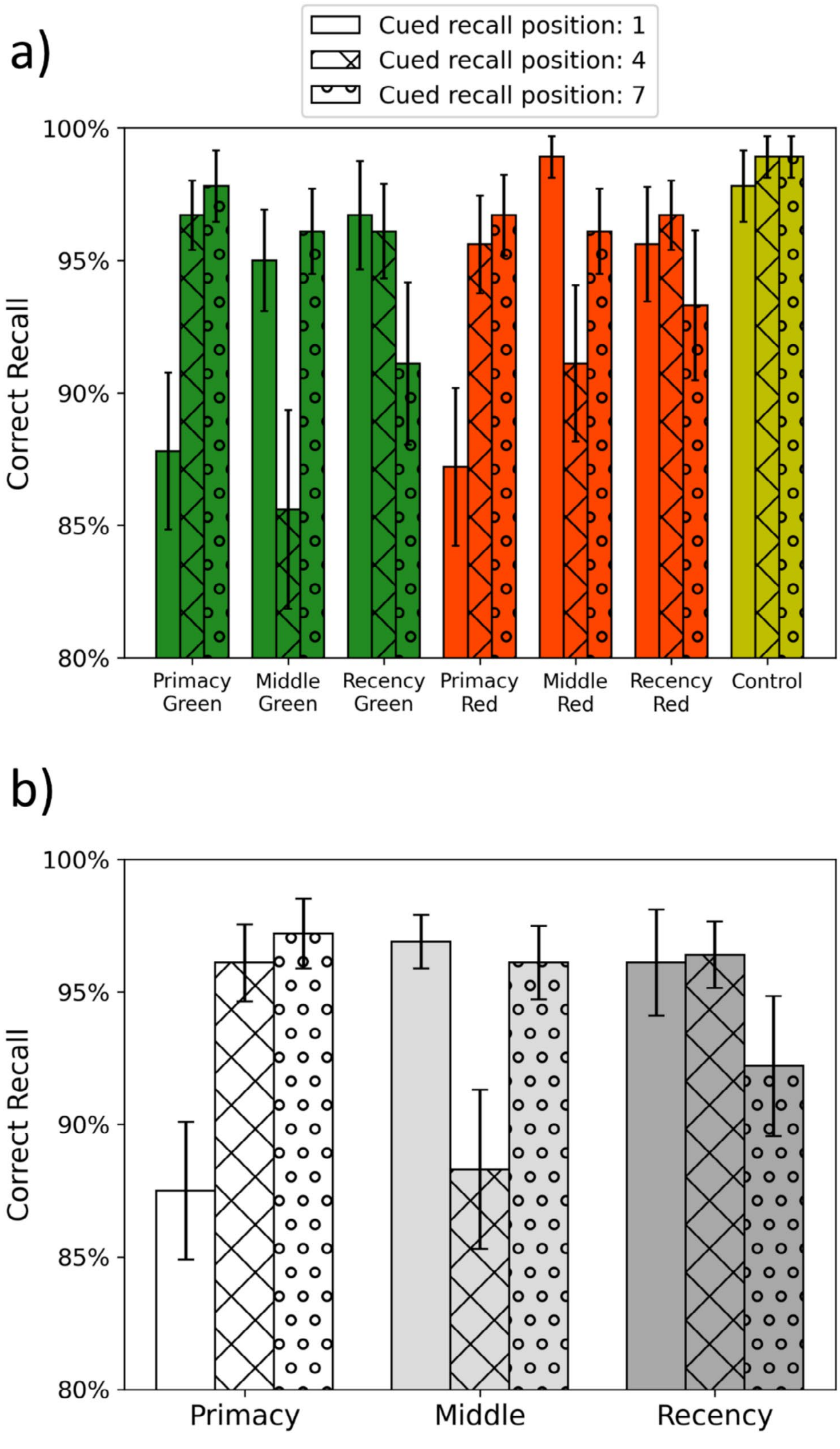

FIGURE 6 | Mean recall accuracy in Experiment 3. The sequences could include a red (environmentally harmful) or a green (environmentally friendly) item in the beginning (primacy), in the middle, or at the end (recency). Control sequences comprised yellow (intermediate environmental footprint) items only. Cues for recall were either drawn from the first (1), middle (4) or last position (7) of the sequence. Panel A shows data separated across sequence types, whereas Panel B shows the means collapsed across sequence types, without the control sequence. Error bars represent standard error of means.

between the cued recall position and sequence type, $F = 9.50$, $p < 0.001$, $\eta_p^2 = 0.139$. Focusing on any recency effects, post hoc Wilcoxon signed-rank tests were used to compare the primacy sequence when the cued recall position was 1, and the recency sequence when the cued recall position was 7, revealing a significant difference, $W = 86$, $p = 0.019$, $BF_{10} = 3.231$ (BF was estimated using the Bayes Wilcoxon signed-rank test). In a similar comparison across the middle (cued recall 4) and recency (cued recall 7) sequences, $W = 23.5$, $p = 0.021$, $BF_{10} = 1.248$, the outcome was also significant. There was no difference across the primacy (cued recall 1) and middle (cued recall 4) sequences, $W = 198$, $p = 0.475$, $BF_{10} = 0.202$. Overall, we observed a recency effect for accuracy in the source memory task, but no primacy effect (Figure 6B). We judged our sample size estimation of $N = 60$ (based on the effect size reported in Talmi and Goshen-Gottstein 2006) to be appropriate for detecting the recency effect in the source memory task.

Taken together, the fact that we observe recency biases for both the rating and source memory tasks suggests that the reason why participants were more impacted by recently presented green or red items in the rating task could be that they are more likely to remember them. This point is central to the current study and will be explored in more detail in Experiment 4.

5 | Experiment 4: Memory of To-Be-Estimated Items After a Filled Retention Interval

A limitation of Experiment 3 was that source memory % correct was very high (Figure 6). This did not allow a direct comparison between ratings in correct and incorrect source memory trials, as there were too few incorrect trials. Experiment 4 seeks to further study the potential interaction between correct recall and ratings of environmental friendliness by introducing a filled retention interval task (as in Experiment 2) that might serve to reduce the recency effect in source memory performance and the rating task. If the recency effect emerges because judgments are based on a mnemonic record of the list items, then the recency effect should be smaller if memory of recently presented items is impaired by the filled retention interval. Experiment 3 provided some evidence for this assumption, but Experiment 4 sought to put the hypothesis under a more critical test. Experiment 4 also offered another opportunity to test the assumption that temporal distinctiveness is the mechanism responsible for the recency effect. If temporal distinctiveness rather than interference or decay is responsible for the recency effect, then the task-filled retention interval should selectively impair recall of recently presented items.

5.1 | Method

5.1.1 | Participants

We aimed for a sample size equal to that of Experiments 2 and 3. Sixty participants (32 female and 27 male and 1 who did not reveal their gender, Mean age = 42.72 years, SD = 11.5) took part in Experiment 4. All participants were English native speakers and were based in the UK; they also reported normal or corrected-to-normal vision and reported no form of colour-blindness. Informed consent was obtained from all participants who were recruited via Prolific and were compensated with £6.75 for participating.

5.1.2 | Materials

The materials used in this experiment were identical to those described in Experiment 3.

5.1.3 | Design and Procedure

The design and procedure of Experiment 4 was identical to Experiment 3, with the addition of a filled retention interval task that participants completed after the shopping sequence (Figure 3; Phase 1) and before the rating task (Figure 3; Phase 2). This retention interval task was the same arithmetic task as the one described in Experiment 2. The inclusion of this retention interval task made the experiment slightly longer, and participants completed the experiment in around 45 min.

5.2 | Results and Discussion

Figure 7 shows the results of the rating task. A repeated measures 2 (colour: with green item vs. with red item) × 3 (sequence type: primacy [red/green item first] vs. middle [red/green item in the middle] vs. recency [red/green item last in the sequence]) analysis of variance with ratings of environmental friendliness as the dependent variable was used to investigate the results of the rating task. In this analysis, the assumption of sphericity was violated only for the sequence factor (Mauchly's W = 0.88), and the rating data for most sequence types violated the assumption of normality (Shapiro–Wilk's test). We report Bayesian factors below but suggest caution in their interpretation. The analysis revealed a significant effect of colour, $F = 67.39$, $p < 0.001$, $\eta_p^2 = 0.533$, $BF_{10} = 1.088 \times 10^8$, no significant effect of sequence type, $F = 0.02$, $p = 0.985$, $\eta_p^2 = 0.000$, $BF_{10} = 0.044$, and no significant interaction between colour and sequence type, $F = 2.92$, $p = 0.058$, $\eta_p^2 = 0.047$, $BF_{10} = 0.111$. This outcome differs from Experiment 3, where the important interaction between colour and sequence type was significant. The Bayesian factors also suggest that sequence type was significant in Experiment 3 but not Experiment 4. These results suggest that the inclusion of the arithmetic retention interval task has a similar effect here as it did in Experiment 2, where it attenuates the recency effect and disrupts memory of the sequences. In the following section, this will be tested more thoroughly in a cross-experiment analysis of Experiments 3 and 4.

We also examined accuracy in the source memory task to further explore effects on memory, focusing on key comparisons across sequence types and cued recall positions. Figure 8 shows accuracy for the different sequence types and cued recall positions, collapsed across the colour factor. Wilcoxon signed-rank tests (including Bayes Wilcoxon signed-rank test) were used to compare the primacy sequence when the cued recall position was 1, and the recency sequence when the cued recall position was 7, revealing no significant difference, $W = 140$, $p = 0.144$, $BF_{10} = 0.400$. In a similar comparison

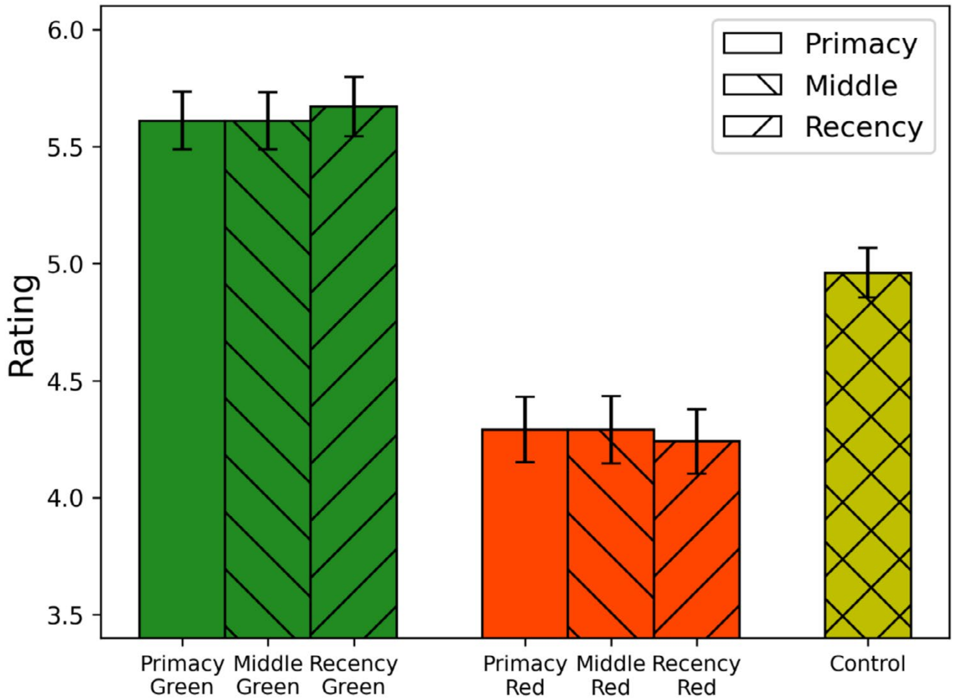

FIGURE 7 | Mean estimates of the environmental friendliness of shopping sequences in Experiment 4. Data is structured like Figure 4. Error bars represent standard error of means.

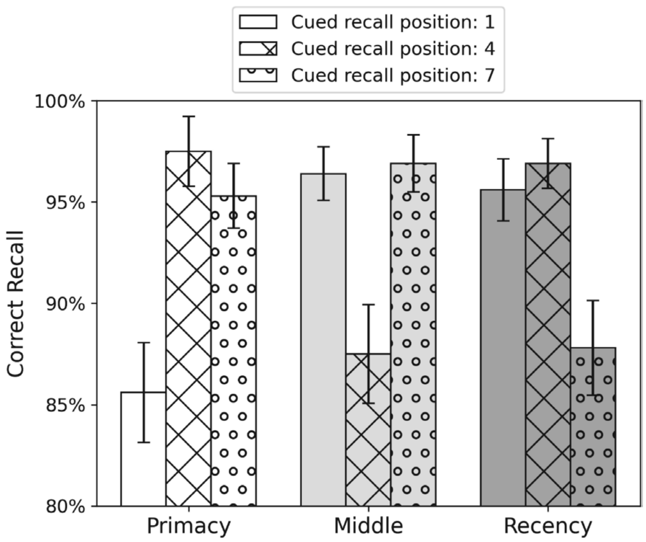

FIGURE 8 | Mean recall accuracy in Experiment 4. Data is structured like Figure 6, Panel B. Error bars represent standard error of means.

across the middle (cued recall 4) and recency (cued recall 7) sequences, $W = 191$, $p = 0.790$, $BF_{10} = 0.144$, the outcome was also not significant. Finally, there was no difference across the primacy (cued recall 1) and middle (cued recall 4) sequences, $W = 199$, $p = 0.329$, $BF_{10} = 0.231$. These key comparisons did not produce any differences, and no indication of a recency effect. This contrasts with the outcome of Experiment 3, where a recency effect was discovered (c.f. Figures 6B and 8). Overall, this informal comparison across experiments suggests that the arithmetic retention interval task has reduced the mnemonic record of the sequences, and this in turn is a likely explanation for the recency effect in judgments of environmental impact. Below follows a more formal comparison of differences across Experiments 3 and 4.

6 | Cross-Experiment Analysis of Experiments 3 and 4

This cross-experiment analysis explores differences in the results from Experiments 3 (recall group only) and 4 in (1) the rating task and (2) the source memory task, where the inclusion of the retention interval task might have impacted judgments of environmental impact and directly affected the memory of the presented items. To analyse differences across the rating task, a repeated measures 2 (colour: with green item vs. with red item) × 3 (sequence type: primacy [red/green item first] vs. middle [red/green item in the middle] vs. recency [red/green item last in the sequence]) × 2 (Experiment: Experiment 3 [recall group] vs. Experiment 4) analysis of variance with ratings of environmental friendliness as the dependent variable was used to investigate the results. In this analysis, the assumption of sphericity was violated only for the sequence factor (but Mauchly's W was high [0.94]), and the data violated the assumption of normality (Shapiro–Wilk's test). The analysis revealed a significant effect of colour, $F = 188.86$, $p < 0.001$, $\eta_p^2 = 0.615$, $BF_{10} = 1.608 \times 10^{14}$, no significant effect of sequence type, $F = 1.21$, $p = 0.300$, $\eta_p^2 = 0.010$, $BF_{10} = 11.788$ (but the Bayes factor favors the hypothesis over the null), and a significant interaction between colour and sequence type, $F = 8.92$, $p < 0.001$, $\eta_p^2 = 0.070$, $BF_{10} = 67.599$. There were no significant interactions between colour and Experiment, or sequence and Experiment. Most importantly, the three-way interaction between colour, sequence, and Experiment was not significant, $F = 1.35$, $p = 0.261$, $\eta_p^2 = 0.011$, $BF_{10} = 5.212 \times 10^{-6}$. While Experiments 3 and 4, when analysed individually, yielded different outcomes regarding recency effects in the rating task (where Experiment 3 showed a significant interaction between colour and sequence type but Experiment 4 did not), the non-significance of this three-way interaction suggests that this difference is not robust.

Regarding accuracy measures in the source memory task, we were primarily interested in three key comparisons across the two experiments. The first comparison (Mann–Whitney U test) is % correct responses for primacy sequences when the cued recall position was 1, $U = 1691$, $p = 0.530$, $BF_{10} = 0.231$ (Bayesian Mann–Whitney U test), indicating no differences across experiments. A similar outcome was seen for middle sequences when the cued recall position was 4, $U = 1603$, $p = 0.230$, $BF_{10} = 0.265$. But, for the most interesting comparison, recency sequences when the cued recall position was 7, we observed a significant difference across experiments, $U = 1389$, $p = 0.008$, $BF_{10} = 0.711$ (but the Bayes factor does not support the hypothesis over the null). This result suggests that source memory was impaired by the retention interval task, attenuating the recency effect in both source memory and ratings.

Overall, the results show a tendency for the recency effect to be attenuated in both the rating task (cf. Figures 4 and 7) and the source memory task (cf. Figures 6B and 8). However, the lack of a significant three-way interaction in the rating task limits our theoretical explanation where we hypothesised a mnemonic account of the recency effect. Yet this account is supported by the fact that recency sequences were remembered better without the retention interval task (cf. Figures 6B and 8).

7 | General Discussion

The present study set out to clarify why retrospective judgments of environmental impact are shaped by a recency effect—greater influence of recently presented items—but not by a corresponding primacy effect. Across four experiments, we found converging evidence that this asymmetry arises from memory processes: recently presented items are more accessible in memory and therefore more influential. Critically, the recency effect diminished when memory for recent items was impaired (e.g., via a filled retention interval), supporting the view that it arises from better memory for recently presented items—possible due to temporal distinctiveness rather than interference or decay. This pattern contrasts with the well-established primacy effect seen in list recall, where the earliest items are typically best remembered. As such, our results suggest functional dissimilarities between retrospective judgments of environmental impact and classical memory phenomena like serial position effects in free or serial recall. The findings also point to broader implications for behaviour change in sustainability contexts, especially where people's evaluations of past eco-behaviour influence future decisions.

7.1 | Theoretical Implications

Our findings demonstrate that retrospective eco-judgments are shaped—at least in part—by the memory accessibility of sequence items at the time of judgment. In support of this view, a filled retention interval reduced the magnitude of the recency effect (Experiments 2 and 4), because the recency effect is a result of superior memory of recently presented items (Experiment 3) and is reduced when memory for recently presented items is reduced (Experiment 4). These findings align with previous research that has shown impaired memory and a reduced recency effect following a filled retention interval (Aldrovandi et al. 2015). The findings are also consistent with the broader role of episodic sampling in judgments (Mason et al. 2023; Moser 1992). We also note that the effect of a filled retention interval on eco-judgments was similar across the two versions of the item encoding task: active decision between pairwise presented items (Experiment 2) and passive viewing of items at presentation (Experiment 4). The purpose of using the pairwise presentation with decision task in Experiments 1 and 2 was to simulate decisions that consumers make during a shopping sequence, while the methodological change to passive viewing was chosen for Experiments 3 and 4 to obtain better experimental control over the involved cognitive processes. By comparing the effect of a filled retention interval on the recency effect in Figures 2 and 7, we conclude that the effect of the filled retention interval was seemingly as large across the more active and passive versions of the encoding task. This suggests that a more active task, which probably leads to deeper encoding processing, does not make people less vulnerable to the effects of a filled retention interval in comparison with more shallow encoding.

The results are also consistent with a temporal distinctiveness explanation of the recency effect, and results are more difficult to reconcile with an interference or decay explanation. Three key findings support a temporal distinctiveness account over decay or interference theories: (a) the recency effect was just as large in magnitude when a concurrent distraction task was performed during item presentation as when the task was not performed (Experiment 1), (b) but smaller when judgments were delayed by a filled retention interval (Experiments 2 and 4), and (c) the filled retention interval impaired memory of recently presented items selectively (Experiment 4). As the relative difference between the first list item and the last list item is the same when continual distraction is performed versus when it is not performed, the recency effect should be similar in magnitude across the two conditions if temporal distinctiveness is the responsible mechanism. A filled retention interval, on the other hand, should have a relatively larger effect on recently presented items. In turn, if decay or interference was the main mechanism responsible for the recency effect, then the retention interval should increase or at least leave the magnitude of the recency effect unchanged, because decay and interference should operate largely on early presented items. Of note, the introduction of a continual distraction following the last item in Experiment 1 also delays judgment. With this relatively short delay, the recency effect remained intact, while the longer delays of Experiments 2 and 4 diminished the effect. This is expected, as items become progressively less temporally distinct with increasing delay. Taken together, the results are consistent with the former alternative and align with previous research suggesting that temporal distinctiveness is responsible for the recency effect in retrospective judgment (Sörqvist et al. 2024b). To our knowledge, this is among the first studies to directly link temporal memory mechanisms to retrospective environmental judgments, bridging the literature on memory and sustainability decision-making.

7.2 | Suggestions for Future Research

In relation to recency effects, the results across all experiments are generally clear and consistent with the temporal distinctiveness view (e.g., Brown et al. 2007). However, here as elsewhere (Aldrovandi et al. 2015; Sörqvist et al. 2024a, 2024b), primacy effects were not found in any of the experiments (e.g., estimates of green primacy lists were consistently lower than green middle lists, and estimates of red primacy lists were consistently higher than red middle lists, in Experiments 3–4, contrary to what would be the case if there was a primacy effect). Some related work (Montgomery and Unnava 2009) shows that the introduction of a delay between learning and test *increases* the influence of earlier presented items on retrospective evaluation. In the current context, this would translate to a more pronounced primacy effect from environmentally significant items following a task-filled retention interval. However, this pattern was not observed in Aldrovandi et al. (2015) and not in Experiment 2 (including only sequences with environmentally significant items first or last) nor in Experiment 4 (also including sequences with environmentally significant items in the middle, which allows for a more direct measure of the possible presence of a primacy effect) of the current series. If anything, the filled retention interval of Experiment 2 *reduced* the influence from environmentally significant initial items. It is unclear how this meshes with a temporal distinctiveness account and presents a challenge for future work.

The evidence for poorer recall of recently presented items following a filled retention interval in Experiment 4 was not statistically strong. A reason for this could have been that the recall task was relatively easy, even with a filled retention interval, as reflected by the high accuracy scores (Figures 6B and 8) that were well above what would be expected from guessing (i.e., an accuracy score of 33.3% correct). Since sequences with green/red items always contained just one green/red item, the rest being yellow, a strategy would be to maintain the only green/red item in the sequence in memory and discard the rest, as this would simplify the memory task. Yet, even under these circumstances, the results indicate a correspondence between memory and judgment such that the recency effect is a result of better memory for recently presented items (Experiment 3) and becomes lower in magnitude when memory of recently presented items is impaired (Experiment 4). Future research should aim to disentangle genuine source memory from guessing by developing more challenging recall tasks and trial structures that yield a wider distribution of accuracy. This would allow for stronger tests of the proposed link between memory and retrospective eco-judgments, which is important as the evidence for a relation between memory performance and retrospective judgments seen here is merely correlational. One possibility is that the experimental manipulations of the retention interval influenced both memory and judgment without a causal link between them. A further exploration of the causal relationship is therefore necessary to draw a firm conclusion about their interdependence.

7.3 | Applied Implications

An overarching aim of the present study was to contribute to the discussion of the role of memory and learning in climate action and sustainability (Zhao et al. 2024) by elucidating how memory processes underpin the perceived environmental friendliness of past events and behaviours. Understanding the role of memory in these processes may help explain why past environmentally friendly behaviuor may sometimes increase the likelihood of behaving less sustainably in the future (Gholamzadehmir et al. 2019; Sörqvist et al. 2024a). If such behavioural spillover effects are viewed as a memory problem—where people must recall what they have done in the past in order for it to influence future behaviour—then understanding these memory processes might help in finding ways to prevent negative spillover and large-scale rebound effects in carbon reduction strategies (Brockway et al. 2021).

A take home message from the present study is that longer periods between encoding and judgment diminish the recency effect in eco-judgments. Furthermore, periods of delay make the perceived environmental friendliness of the past regress toward an average. More environmentally friendly sequences are perceived as less environmentally friendly after the delay, and vice versa. This can have implications for future decision-making also in settings more representative of the real-world (e.g., web-shop interfaces). For example, e-commerce platforms might mitigate undesirable spillover by introducing a delay between purchase and delivery choice, thereby reducing the mnemonic influence of recently selected products. Future research could address memory processes involved in the build-up of self-perception from environmentally friendly and environmentally harmful behaviour over the long term and how this self-perception determines environment-related decision-making.

The current paper is primarily concerned with order effect from manipulations of the temporal position of list items. It should be noted though that the perceived environmental friendliness of sequences depended to a much greater extent on *what* was presented rather than *when* it was presented. Thus, item-identity seems to have a greater potential than item-order at producing spillover effects on post-sequence decision-making. Furthermore, the sequences of the current study involved either a single instance of an environmentally significant item or no environmentally significant item at all. When people buy articles at a shopping mall or a web-shop, for example, the shopping sequence will be of varying quantity and often comprise articles with differing environmental impacts. It is possible that order effects interact with item quantity, such that the selection of a large quantity of environmentally significant items at the end of a shopping sequence produces a disproportionately strong recency effect on perceived environmental friendliness of the shopping sequence. Future research should explore ways by which sequences with a mixture of items and quantities interact with order effects to produce spillover effects on subsequent decision-making. Understanding these memory processes is key to understanding how people build a view of the environmental consequences of their own and others' actions, with implications for both immediate and long-term environmentally significant behaviour. By revealing how memory dynamics shape perceived environmental impact, this work offers a foundation for behaviourally informed strategies aimed at reducing rebound effects and promoting sustainable long-term decision-making.

Author Contributions

Patrik Sörqvist: conceptualization, methodology, formal analysis, funding acquisition, writing – original draft, supervision, project administration. **Emil Skog:** conceptualization, methodology, software, data curation, investigation, formal analysis, visualization, writing – original draft. **Johanna Heidenreich:** investigation, software, methodology. **John E. Marsh:** conceptualization, writing – original draft, writing – review and editing.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Open Science Framework at https://doi.org/10.17605/OSF.IO/EXHAV, reference number 10.17605/OSF.IO/EXHAV.

References

Alaybek, B., R. S. Dalal, S. Fyffe, et al. 2022. "All's Well That Ends (And Peaks) Well? A Meta-Analysis of the Peak-End Rule and Duration Neglect." *Organizational Behavior and Human Decision Processes* 170: 104149. https://doi.org/10.1016/j.obhdp.2022.104149.

Aldrovandi, S., M. Poirier, P. Kusev, and P. Ayton. 2015. "Retrospective Evaluations of Sequences: Testing the Predictions of a Memory-Based Analysis." *Experimental Psychology* 62: 320–334. https://doi.org/10.1027/1618-3169/a000301.

Asutay, E., A. Genevsky, L. F. Barrett, J. P. Hamilton, P. Slovic, and D. Västfjäll. 2021. "Affective Calculus: The Construction of Affect Through Information Integration Over Time." *Emotion* 21: 159–174. https://doi.org/10.1037/emo0000681.

Bjork, R. A., and W. B. Whitten. 1974. "Recency-Sensitive Retrieval Processes in Long-Term Free Recall." *Cognitive Psychology* 6: 173–189. https://doi.org/10.1016/0010-0285(74)90009-7.

Bornstein, B. H., C. B. Neely, and D. C. LeCompte. 1995. "Visual Distinctiveness Can Enhance the Recency Effect." *Memory and Cognition* 23: 273–278. https://doi.org/10.3758/bf03197229.

Brockway, P. E., S. Sorrell, G. Semieniuk, M. K. Heun, and V. Court. 2021. "Energy Efficiency and Economy-Wide Rebound Effects: A Review of the Evidence and Its Implications." *Renewable and Sustainable Energy Reviews* 141: 110781. https://doi.org/10.1016/j.rser.2021.110781.

Brown, G. D., I. Neath, and N. Chater. 2007. "A Temporal Ratio Model of Memory." *Psychological Review* 114: 539–576. https://doi.org/10.1037/0033-295X.114.3.539.

Brown, J. 1958. "Some Tests of the Decay Theory of Immediate Memory." *Quarterly Journal of Experimental Psychology* 10: 12–21. https://doi.org/10.1080/174702158084162.

Chajut, E., A. Caspi, R. Chen, M. Hod, and D. Ariely. 2014. "In Pain Thou Shalt Bring Forth Children: The Peak-And-End Rule in Recall of Labor Pain." *Psychological Science* 25: 2266–2271. https://doi.org/10.1177/0956797614551004.

Collins, D. J., and D. R. Shanks. 2002. "Momentary and Integrative Response Strategies in Causal Judgment." *Memory & Cognition* 30: 1138–1147. https://doi.org/10.3758/BF03194331.

Cowan, N., J. S. Saults, E. M. Elliott, and M. V. Moreno. 2002. "Deconfounding Serial Recall." *Journal of Memory and Language* 46: 153–177. https://doi.org/10.1006/jmla.2001.2805.

Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39: 175–191. https://doi.org/10.3758/bf03193146.

Fredrickson, B. L. 2000. "Extracting Meaning From Past Affective Experiences: The Importance of Peaks, Ends, and Specific Emotions." *Cognition and Emotion* 14: 577–606. https://doi.org/10.1080/026999300402808.

Fredrickson, B. L., and D. Kahneman. 1993. "Duration Neglect in Retrospective Evaluations of Affective Episodes." *Journal of Personality and Social Psychology* 65: 45–55. https://doi.org/10.1037/0022-3514.65.1.45.

Gholamzadehmir, M., P. Sparks, and T. Farsides. 2019. "Moral Licensing, Moral Cleansing and Pro-Environmental Behavior: The Moderating Role of Pro-Environmental Attitudes." *Journal of Environmental Psychology* 65: 101334. https://doi.org/10.1016/j.jenvp.2019.101334.

Glanzer, M., and A. R. Cunitz. 1966. "Two Storage Mechanisms in Free Recall." *Journal of Verbal Learning and Verbal Behavior* 5: 351–360. https://doi.org/10.1016/S0022-5371(66)80044-0.

Glenberg, A. M., M. M. Bradley, T. A. Kraus, and G. J. Renzaglia. 1983. "Studies of the Long-Term Recency Effect: Support for a Contextually Guided Retrieval Hypothesis." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9: 231–255. https://doi.org/10.1037/0278-7393.9.2.231.

Howard, M. W., M. S. Fotedar, A. V. Datey, and M. E. Hasselmo. 2005. "The Temporal Context Model in Spatial Navigation and Relational Learning: Toward a Common Explanation of Medial Temporal Lobe Function Across Domains." *Psychological Review* 112: 75–116. https://doi.org/10.1037/0033-295X.112.1.75.

Howard, M. W., and M. J. Kahana. 1999. "Contextual Variability and Serial Position Effects in Free Recall." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25: 923–941. https://doi.org/10.1037//0278-7393.25.4.923.

Kahneman, D., B. L. Fredrickson, C. A. Schreiber, and D. A. Redelmeier. 1993. "When More Pain Is Preferred to Less: Adding a Better End." *Psychological Science* 4: 401–405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x.

Kemp, S., C. D. Burt, and L. Furneaux. 2008. "A Test of the Peak-End Rule With Extended Autobiographical Events." *Memory & Cognition* 36: 132–138. https://doi.org/10.3758/mc.36.1.132.

Lakens, D. 2013. "Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-Tests and ANOVAs." *Frontiers in Psychology* 4: 863. https://doi.org/10.3389/fpsyg.2013.00863.

Lee, M. D., and E. J. Wagenmakers. 2014. *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Manning, S. K., and H. Schreier. 1988. "Recency and Suffix Effects in Pictures as a Function of Recall Method." *American Journal of Psychology* 101: 97–109. https://doi.org/10.2307/1422796.

Mantonakis, A., P. Rodero, I. Lesschaevem, and R. Hasti. 2009. "Order in Choice: Effects of Serial Position on Preferences." *Psychological Science* 20: 1309–1312. https://doi.org/10.1111/j.1467-9280.2009.02453.x.

Mason, A., G. D. A. Brown, G. Ward, and S. Farrell. 2023. "The Role of Episodic Memory Sampling in Evaluation." *Psychonomic Bulletin & Review* 31: 1353–1363. https://doi.org/10.3758/s13423-023-02413-z.

Montgomery, N. V., and H. R. Unnava. 2009. "Temporal Sequence Effects: A Memory Framework." *Journal of Consumer Research* 36: 83–92. https://doi.org/10.1086/595278.

Moser, D. V. 1992. "Does Memory Affect Judgment? Self-Generated Versus Recall Memory Measures." *Journal of Personality and Social Psychology* 62: 555–563. https://doi.org/10.1037//0022-3514.62.4.555.

Oberauer, K., S. Lewandowsky, S. Farrell, C. Jarrold, and M. Greaves. 2012. "Modeling Working Memory: An Interference Model of Complex

Span." *Psychonomic Bulletin and Review* 19: 779–819. https://doi.org/10.3758/s13423-012-0272-4.

Palan, S., and C. Schitter. 2018. "Prolific.Ac—A Subject Pool for Online Experiments." *Journal of Behavioral and Experimental Finance* 17: 22–27. https://doi.org/10.1016/j.jbef.2017.12.004.

Peirce, J., J. R. Gray, S. Simpson, et al. 2019. "PsychoPy2: Experiments in Behavior Made Easy." *Behavior Research Methods* 51, no. 1: 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Penny, C. G. 1989. "Modality Effects and the Structure of Short-Term Verbal Memory." *Memory and Cognition* 17: 398–422. https://doi.org/10.3758/BF03202613.

Peterson, L., and M. J. Peterson. 1959. "Short-Term Retention of Individual Verbal Items." *Journal of Experimental Psychology* 58: 193–198. https://doi.org/10.1037/h0049234.

Postman, L., and L. W. Phillips. 1965. "Short-Term Temporal Changes in Free Recall." *Quarterly Journal of Experimental Psychology* 17: 132–138. https://doi.org/10.1080/17470216508416422.

Redelmeier, D. A., and D. Kahneman. 1996. "Patients' Memories of Painful Medical Treatments: Real-Time and Retrospective Evaluations of Two Minimally Invasive Procedures." *Pain* 66: 3–8. https://doi.org/10.1016/0304-3959(96)02994-6.

Rey, A., K. Le Goff, M. Abadie, and P. Courrieu. 2019. "The Primacy Order Effect in Complex Decision Making." *Psychological Research* 84: 1739–1748. https://doi.org/10.1007/s00426-019-01178-2.

Robinson, E., J. Blissett, and S. Higgs. 2013. "The Influence of Recent Tasting Experience on Expected Liking for Foods." *Food Quality and Preference* 27, no. 1: 101–106. https://doi.org/10.1016/j.foodqual.2012.07.001.

Schreiber, C. A., and D. Kahneman. 2000. "Determinants of the Remembered Utility of Aversive Sounds." *Journal of Experimental Psychology: General* 129: 27–42. https://doi.org/10.1037//0096-3445.129.1.27.

Sörqvist, P., E. Skog, J. Heidenreich, and J. E. Marsh. 2024c. "All's Eco-Friendly That Ends Eco-Friend (If Remembered as Such): Memory Processes in Retrospective Judgment of Environmentally Significant Sequences." Dataset available at *Open Science Framework*. https://doi.org/10.17605/OSF.IO/EXHAV.

Sörqvist, P., J. Heidenreich, B. Hoxha, H. Johansson, and J. E. Marsh. 2024a. "Short-Term Memory Effects of Eco-Labeling: Evidence From the Perceived Environmental Friendliness of Sequential Consumer Behavior." *Food Quality and Preference* 121: 105276.

Sörqvist, P., S. Lindeberg, and J. E. Marsh. 2024b. "All's Eco-Friendly That Ends Eco-Friendly: Short-Term Memory Effects in Carbon Footprint Estimates of Temporal Item Sequences." *Applied Cognitive Psychology* 38: e4204.

Talmi, D., and Y. Goshen-Gottstein. 2006. "The Long-Term Recency Effect in Recognition Memory." *Memory* 14: 424–436. https://doi.org/10.1080/09658210500426623.

Tan, L., and G. Ward. 2000. "A Recency-Based Account of the Primacy Effect in Free Recall." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 1589–1625. https://doi.org/10.1037//0278-7393.26.6.1589.

Tan, L., G. Ward, and R. Grenfell-Essam. 2010. "Examining the Relationship Between Free Recall and Immediate Serial Recall: The Effects of List Length and Output Order." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36: 1207–1241. https://doi.org/10.1037/a0020122.

Thapar, A., and R. L. Greene. 1993. "Evidence Against a Short-Term-Store Account of Long-Term Recency Effects." *Memory and Cognition* 21: 329–337. https://doi.org/10.3758/BF03208265.

Tzeng, O. J. 1973. "Positive Recency Effect in a Delayed Free Recall." *Journal of Verbal Learning and Verbal Behavior* 12, no. 4: 436–439. https://doi.org/10.1016/S0022-5371(73)80023-4.

van den Bergh, D., E.-J. Wagenmakers, and F. Aust. 2023. "Bayesian Repeated-Measures Analysis of Variance: An Updated Methodology Implemented in JASP." *Advances in Methods and Practices in Psychological Science* 6: 1–11. https://doi.org/10.1177/25152459231168024.

Ward, G. 2002. "A Recency-Based Account of the List Length Effect in Free Recall." *Memory and Cognition* 30: 885–892. https://doi.org/10.3758/bf03195774.

Watkins, M. J., I. Neath, and E. S. Sechler. 1989. "Recency Effect in Recall of a Word List When an Immediate Memory Task Is Performed After Each Word Presentation." *American Journal of Psychology* 102: 265–270. https://doi.org/10.2307/1422957.

Xu, L., M. Niculescu, and M. Hyman. 2023. "When Interruptions Can Boost Sales: An On-Line Versus Memory-Based Perspective." *International Journal of Consumer Studies* 47: 1733–1748. https://doi.org/10.1111/ijcs.12941.

Zare, S., and P. Pearce. 2018. "Order Effects and Multi-City Visits: Tour Guides' Perspectives." *International Journal of Tourism Cities* 4: 194–206. https://doi.org/10.1108/IJTC-08-2017-0042.

Zare, S., and P. L. Pearce. 2022. "Does the Order of Visiting Destinations Affect Their Recall and Evaluation?" *Journal of Travel Research* 61: 1559–1572. https://doi.org/10.1177/00472875211042673.

Zhao, J., J. Radke, F. S. Chen, S. Sachdeva, S. J. Gershman, and Y. Luo. 2024. "How Do We Reinforce Climate Action?" *Sustainability Science* 19: 1503–1517. https://doi.org/10.1007/s11625-024-01486-6.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** acp70103-sup-0001-Supinfo.docx.