Journal of Mathematical Economics 109 (2023) 102913

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Context-sensitive rationality: Choice by salience☆

Alfio Giarlotta [a], Angelo Petralia [a,*], Stephen Watson [b]

[a] Department of Economics and Business, University of Catania, Italy
[b] Department of Mathematics and Statistics, York University, Canada

ARTICLE INFO

Manuscript handled by Editor H. Sean

Keywords:
Bounded rationality
Context-sensitive rationalization
Salience
Moody choice
Rationalization by multiple rationales
Limited attention

ABSTRACT

We describe a context-sensitive approach to individual choice, in which the explanation is provided by a family of linear orders indexed by all available items. Selection from a menu is then recovered by the classical maximization paradigm, subject to the constraint that the justifying rationale must be indexed by an item of the menu. This approach allows us to pursue two complementary goals: (1) a fine classification of all possible choices into classes of rationality, and (2) a bounded rationality model based on an ordinal notion of salience. Concerning (1), we refine the context-free model of *rationalization by multiple rationales*, partitioning the class of all choice functions on n items into n classes of rationality. The least rational class is expressive of a moody behavior, which is rare for small n, but prevailing for large n. Concerning (2), we enrich our framework by a binary relation of salience, which guides the selection process. Upon requiring that all rationales associated to equally salient items coincide, choice is explained by appealing to the unique linear order indexed by a maximally salient item of the menu. Choice by salience is a specification of *choice with limited attention*. Numerical estimates show the sharp selectivity of this model of bounded rationality.

Introduction

In this paper we describe a context-sensitive approach to individual choice, in which the informative content of available alternatives forges the judgement of a decision maker (DM). Our main assumption is that each alternative can be looked at from two complementary points of view: (1) *informativeness*, which synthesizes the clues provided by the item about the choice context; (2) *attractiveness*, which is related to the possibility that the item be selected by the DM. Typically, these two aspects are unrelated: for instance, the item frog's legs in a restaurant menu may be totally unattractive to me (and so I will never select it), and yet it catches my attention, delivering important information about the chef's skills (and so convincing me to order a dish that I would otherwise avoid).

The explanation to choice behavior is here provided by a family of rationales (linear orders). Our approach is context-sensitive in the sense that each rationale is attached to – that is, indexed by – an available item. Therefore, there are as many rationales as elements in the ground set. (Note that, however, these rationales need not be pairwise distinct.) Selection from a menu is then justified by maximizing a linear order *indexed by an item in the menu* itself. This simple constraint allows one to take into account both the informativeness and the attractiveness of alternatives. Specifically, the informativeness of an item suggests a linear order to apply in the decision process, whereas its attractiveness is described by the position it occupies in each rationale.

The effects of informativeness on judgements are first documented by Taylor and Fiske (1978), who, rephrasing Tversky and Kahneman (1974), assert that *"Instead of reviewing all the evidence that bears upon a particular problem, people frequently use the information which is most salient or available to them"*. More recently, the role of context-delivered information on individual choices has been formalized, for instance, by Salant and Rubinstein (2006), Goldin and Reck (2020), and Bhattachary et al. (2021). These authors incorporate the *framing* of choices

☆ The authors wish to thank the Associate Editor Sean Horan and two anonymous referees for their comments and suggestions, which substantially improved the presentation of the paper. They are also very grateful to Larbi Alaoui, José Apisteguia, Davide Carpentiere, Yi-Chun Chen, Francesco Drago, Paolo Ghirardato, M. Ali Khan, Xi Zhi Lim, Paola Manzini, Marco Mariotti, Daniele Pennesi, John K.-H. Quah, Francesco Reito, Marzena J. Rostek, and Xinyang Wang for several insightful observations. A special thank goes to the participants to the 2022 BRIC meeting in Prague, the 2022 Asian School in Economic Theory of the Econometric Society, the 2023 East and South-East Asian Meeting of the Econometric Society, and the Economic Seminar at the National University of Singapore. Alfio Giarlotta acknowledges the support of "Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), PRIN 2017", project *Multiple Criteria Decision Analysis and Multiple Criteria Decision Theory*, grant 2017CY2NCA. Angelo Petralia acknowledges the support of "Ministero del Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), PE9 GRINS "Spoke 8", project *Growing, Resilient, INclusive, and Sustainable*, CUP E63C22002120006.

* Corresponding author.
E-mail addresses: alfio.giarlotta@unict.it (A. Giarlotta), angelo.petralia@unict.it (A. Petralia), swatson@yorku.ca (S. Watson).

https://doi.org/10.1016/j.jmateco.2023.102913
Received 27 December 2022; Received in revised form 26 August 2023; Accepted 19 October 2023
Available online 30 October 2023

in deterministic and stochastic settings; however, they do not explicitly model how the informative content of items may impact the DM's preference.

We take a different path: we assume that available items may convey relevant information about choice, which in turn affects individual judgement. Formally, given a finite set X of alternatives, a *context-sensitive multi-rationalization* for a choice function on all nonempty subsets of X (menus) consists of a family $\mathscr{L} = \{\rhd_x : x \in X\}$ of linear orders (rationales) on X. Then, the informativeness of alternatives guides choice on each menu $A \subseteq X$ in two sequential stages: (1) the DM's attention is captured by some item x of A; (2) the DM explains the unique selection from A by maximizing the linear order $\rhd_x$ indexed by x.

We use a famous example due to Luce and Raiffa (1957) to illustrate our approach.

Example 1 (*Luce and Raiffa's Dinner*). Thea selects a main course from a restaurant menu. She prefers steak (s) over chicken (c) as long as steak is duly cooked, whereas she is not interested in exotic dishes such as frog's legs (f). Thea chooses chicken over steak when they are the only available items, but selects steak if the item frog's legs appears in the menu. This happens because the presence of frog's legs in the menu is perceived by Thea as a sign that the chef knows how to grill a steak. Formally, the choice function c on $X = \{c, f, s\}$ is defined as follows (the unique item selected from each menu is underlined):

$$c f \underline{s}\,, \quad \underline{c} f\,, \quad \underline{c} s\,, \quad f \underline{s}\,.$$

This choice cannot be justified by maximizing a single binary relation, because it violates Axiom α (Chernoff, 1954; Sen, 1971).

On the other hand, we can provide a natural explanation of Thea's behavior by appealing to a context-sensitive multi-rationalization, which only employs two distinct linear orders. To that end, observe preliminarily that frog's legs is definitively the most informative item of all, whereas steak and chicken are rather ordinary items, which may be considered as having the same (low) informativeness. Moreover, the most attractive items are steak or chicken (depending on the context), whereas frog's legs is not attractive at all.

These observations yield the following family $\mathscr{L} = \{\rhd_c, \rhd_f, \rhd_s\}$ of preferences on X:

$$c \rhd_c s \rhd_c f\,, \quad s \rhd_f c \rhd_f f\,, \quad \rhd_s = \rhd_c\,.$$

Note that: (i) the linear orders indexed by the least informative items (steak and chicken) coincide; (ii) the most attractive items in the two linear orders are steak and chicken, whereas frog's legs is always the least attractive item. Then choice from any menu A is explained by maximizing a linear order in $\mathscr{L}$ indexed by some item in A. For instance, the maximization of $\rhd_f$ justifies the selection of s from X (as it should be, because the presence of frog's legs in the menu induces to choose steak). Similarly, the maximization the linear order $\rhd_s = \rhd_c$ explains the selection of c from $\{s, c\}$.

Luce and Raiffa's dinner is also used by Kalai et al. (2002) to illustrate the paradigm of *rationalization by multiple rationales (RMR)*. According to this approach, the DM may use several rationales (linear orders) to justify her choice: she selects from each menu the unique element that is maximal according to one of these preferences. The authors label their approach *context-free*, because the selection of a rationalizing order from the available ones is not structurally linked to the menu itself. Then the minimum number of linear orders needed in an RMR justification yields a *context-free rationality index* of a choice: the larger this number, the less rational the behavior.

Our context-sensitive paradigm is essentially the same as the RMR approach from the point of view of attractiveness: in fact, it explains choice behavior by the maximization of some of the available rationales. However, it yields a refinement of RMR, insofar as the informativeness of items has a relevant effect in the decision process as well. Indeed, to explain choice behavior in a menu the DM is not allowed to use *any* rationale in the given list, but she is bound to maximize one of those indexed by items *belonging* to the menu.

Our approach yields a *context-sensitive rationality index*, defined – similarly to the RMR context-free rationality index – as the number of distinct rationales required for a context-sensitive justification of the observed behavior. As for the RMR, the derived partition of choices into disjoint classes of rationality provides a discrimination of behaviors according to their internal coherence. We pay special attention to maximally irrational choices, which are those requiring as many distinct rationales as the total number of items: these behaviors are typical of a *moody* DM.

In the RMR model, the least rational class is nonempty even for a tiny number of alternatives: for instance, Luce and Raiffa's dinner belongs to this class for a ground set of size 3, even if the corresponding behavior can hardly be considered irrational. On the contrary, moodiness is rare for a small number of alternative, and only appears for large ground sets (Theorem 1, *moodiness exists*). However – and not surprisingly – the incidence of moodiness becomes increasingly large when the ground set grows big: in fact, we prove that as the total number of items diverges to infinity, the fraction of moody choices tends to one (Theorem 2, *moodiness asymptotically prevails*). This last fact gives empirical value to our context-sensitive classification of all potential types of choice behavior. In fact, while it is often the case that the DM's behavior may be somehow justified if she is presented with a very small number of options, it is experimentally proved that she typically exhibits strong inconsistencies in preferences whenever the ground set of alternatives is large.

The general version of our context-sensitive approach does not reveal the process that yields informative alternatives to shape the DM's preference. In fact, as context-sensitivity explains any type of potential behavior, it can only be used with a purpose of classification, that is, labeling choices according to their index of rationality; in particular, it cannot be elicited from choice data.

However, a meaningful specification of our general framework can be given full empirical content. Formally, upon adding a binary relation of salience that encodes the notability of items, and imposing that rationales indexed by equally salient items are the same, we obtain a testable model of bounded rationality. In the second part of this paper, we study this testable specification of context-sensitive rationality, called *choice by salience*, from several points of view.

The salience order provides an ordinal evaluation of how intriguing an item is in the DM's eyes when compared to a different one. In our model, we impose two conditions: (i) salience is a weak order on X; (ii) all linear orders indexed by salience-indifferent items are equal. Then a salient explanation for the selection from a menu A goes as follows: first the DM exclusively focuses her attention on the most salient items on A; then she selects the alternative maximizing the (unique) linear order associated to them. We use again Luce and Raiffa (1957)'s example to illustrate choice by salience.

Example 2 (*Luce and Raiffa's Dinner, Continued*). Let $\mathscr{L} = \{\rhd_c, \rhd_f, \rhd_s\}$ be the family of linear orders on $X = \{c, f, s\}$ defined as in Example 1. Furthermore, let $\succsim$ be the weak order on X such that $f \succ c, s$, and $c \sim s$. (Here $\succ$ means 'is strictly more salient than', whereas $\sim$ stands for 'has the same salience as'.) Selection from any menu A is explained by maximizing the unique linear order in $\mathscr{L}$ indexed by the most salient items of A. For instance, the most salient item in $X = \{c, f, s\}$ is f, and the maximization of $\rhd_f$ justifies the selection of s from X. Similarly, in $A = \{s, c\}$ the items s and c are equally salient, hence maximizing the linear order $\rhd_s = \rhd_c$ explains the selection of c from A.

Choice by salience is independent from most existing models of bounded rationality, being however a special case of *choice with limited attention* (Masatlioglu et al., 2012). In this paper we provide a characterization of choice by salience, which relies on (i) the asymmetry of a derived relation of revealed salience, and (ii) some properties that the DM's consideration must satisfy when we see the model as a specification of choice with limited attention (Theorem 3, *characterization of choice by salience*). The first characterization is useful for computational reasons, because it is easy to test the asymmetry of a binary relation.[1] The second characterization provides an alternative representation of choice by salience that reproduces the joint effect of limited attention and salience of items on the DM's perception. We conclude the paper by providing some numerical estimates, which show the sharp selectivity of our testable model of bounded rationality (Theorem 4, *selectivity of choice by salience*).

The paper is organized as follows. Section 1 collects some preliminary notions. Section 2 describes the general context-sensitive paradigm, showing that moodiness exists and asymptotically prevails. Section 3 discusses the testable model of choice by salience, providing an algorithmically effective characterization and some numerical estimates. Section 4 collects final remarks and possible directions of research. All proofs are located in the Appendix, with the exception of the very involved proofs of Theorems 1 and 2, which are presented in an Online Appendix.

1. Preliminaries

Here we recall basic notions about choice and preference. Let X be a finite nonempty set of alternatives (*ground set*). Then $\mathscr{X}$ denotes the family of all nonempty subsets of X, and any A in $\mathscr{X}$ is a *menu*. A *choice correspondence* on X is a map $\Gamma : \mathscr{X} \to \mathscr{X}$ that selects at least one item from each menu, that is, $\varnothing \neq \Gamma(A) \subseteq A$ for any $A \in \mathscr{X}$. A *choice function* is a choice correspondence that selects a single item from each menu; thus, we identify it with a map $c : \mathscr{X} \to X$ such that $c(A) \in A$ for any $A \in \mathscr{X}$. We mostly deal with choice functions, and only occasionally with correspondences; thus, we often use 'choice' in place of 'choice function'. Furthermore, to simplify notation we often omit set delimiters and commas: thus, $A \cup x$ stands for $A \cup \{x\}$, $c(xy)$ for $c(\{x, y\})$, etc.

A binary relation R on X is a subset of $X \times X$, which is:
- *reflexive* if xRx, for all $x \in X$;
- *asymmetric* if xRy implies $\neg(yRx)$, for all $x, y \in X$;
- *symmetric* if xRy implies yRx, for all $x, y \in X$;
- *antisymmetric* if xRy and yRx implies $x = y$, for all $x, y \in X$;
- *transitive* if xRy and yRz implies xRz, for all $x, y, z \in X$;
- *acyclic*[2] if $x_1 R x_2 R \dots R x_n R x_1$ holds for no $x_1, x_2, \dots, x_n \in X$, with $n \geqslant 3$;
- *complete* if either xRy or yRx (or both) holds, for all distinct $x, y \in X$.

In this paper we use the symbol $\succsim$ to denote a reflexive binary relation on X, which is interpreted as a *weak preference* on the set of alternatives. The following derived relations are associated to a weak preference $\succsim$ (the elements x, y range over X):
- *strict preference* $\succ$, defined by $x \succ y$ if $x \succsim y$ and $\neg(y \succsim x)$;
- *indifference* $\sim$, defined by $x \sim y$ if $x \succsim y$ and $y \succsim x$;
- *incomparability* $\perp$, defined by $x \perp y$ if $\neg(x \succsim y)$ and $\neg(y \succsim x)$.

Note that $\succ$ is asymmetric and $\sim$ is symmetric, and their union gives $\succsim$.

A weak preference $\succsim$ on X is a *suborder* if $\succ$ is acyclic, a *weak order* if it is transitive and complete, and a *linear order* if it is an antisymmetric weak order. We denote by $\rhd$ the strict part of a linear order (asymmetric, transitive, and complete).

The theory of revealed preferences pioneered by Samuelson (1938) studies when a binary relation suffices to explain choice behavior by maximization. Given a suborder $\succsim$ on X and a menu $A \in \mathscr{X}$, the set of $\succsim$-*maximal* elements of A is[3]

$$\max(A, \succsim) = \{x \in X \,:\, y \succ x \text{ for no } y \in A\} \neq \varnothing\,.$$

A choice $c : \mathscr{X} \to X$ is *rationalizable* if there exists a suborder (in fact, a linear order) $\rhd$ on X such that $c(A) \in \max(A, \rhd)$ for any $A \in \mathscr{X}$. As customary, we abuse notation, and write $c(A) = \max(A, \rhd)$ in place of $c(A) \in \max(A, \rhd)$.

The rationalizability of a choice function[4] is characterized by the property of *contraction consistency* due to Chernoff (1954), also called *Axiom* α by Sen (1971). This property states that if an item is chosen in a menu, then it is also chosen in any submenu containing it:

Chernoff property (Axiom α): for all $A, B \in \mathscr{X}$ and $x \in X$, if $x \in A \subseteq B$ and $c(B) = x$, then $c(A) = x$.

Irrational features of choice behavior are related to violations of Chernoff property.

Definition 1. For any choice $c : \mathscr{X} \to X$, a *switch* is an ordered pair (A, B) of menus such that $A \subseteq B$ and $c(A) \neq c(B) \in A$; in particular, it is *minimal* if $|B \setminus A| = 1$. Equivalently, a minimal switch is a pair $(A, A \cup x)$ of menus such that $c(A) \neq c(A \cup x) \neq x$.

Switches are violations of Axiom α. A minimal switch $(A, A \cup x)$ arises when, once the DM chooses y from A, and a new item x is added to A, then the item selected from the larger menu $A \cup x$ is neither the old nor the new. Due to the finiteness of the ground set X, any switch can be reduced to a minimal one:

Lemma 1. *Let $c : \mathscr{X} \to X$ be a choice. For any switch (A, B), there are a menu $C \in \mathscr{X}$ and an item $x \in X$ such that $A \subseteq C \subsetneq C \cup x \subseteq B$ and $(C, C \cup x)$ is a switch.*

2. Context-sensitive rationality

Here we introduce a general definition of context-sensitive multi-rationalization. This notion yields a classification of all types of choice behavior according to their (context-sensitive) degree of rationality. We conclude this section by analyzing the least rational class of choices, which is expressive of a moody behavior.

2.1. The general paradigm

Definition 2. A *context-sensitive multi-rationalization (CSMR)* of a choice $c : \mathscr{X} \to X$ is a family $\mathscr{L} = \{\rhd_x \,:\, x \in X\}$ of linear orders on X (the *rationales*) such that for any $A \in \mathscr{X}$, the equality $c(A) = \max(A, \rhd_x)$ holds for some $x \in A$.

While the term 'context-sensitive' certainly requires a justification, the term 'multi-rationalization' is self-explanatory: the DM may resort to the possibility of using several rationales (linear orders) instead of a single one in the process of explaining the selection of an item from a menu. This feature is shared with the RMR model:

[1] Some models of bounded rationality require to test the asymmetry *and* the acyclicity of an induced binary relation, which is computationally much more demanding.

[2] Sometimes a binary relation is called *acyclic* if there is no cycle of length $\geqslant 2$ (see. e.g., Masatlioglu et al., 2012): according to this terminology, asymmetry is a special case of acyclicity. We prefer to differentiate the properties of asymmetry and acyclicity, using the former for the absence of cycles of length two, and the latter for the absence of cycles of length at least three.

[3] Note that $\max(A, \succsim) \neq \varnothing$ because X is finite and $\succ$ is acyclic.

[4] For a choice *correspondence*, rationalizability is characterized by Axioms α and γ (Sen, 1971).

Definition 3 (*Kalai et al., 2002*)**.** A *rationalization by multiple rationals (RMR)* of a choice $c : \mathscr{X} \to X$ is a family $\mathscr{L} = \{\rhd_x : x \in X\}$ of linear orders on X (the *rationales*) such that for any $A \in \mathscr{X}$, the equality $c(A) = \max(A, \rhd_x)$ holds for some $x \in X$.[5]

In the RMR model any choice is collectively explained by a collection of rationales, but there is no structural connection between available alternatives and rationalizing linear orders. In other words, attractiveness of items is encoded by several rankings, but their informativeness is disregarded altogether. As a matter of fact, the indexing of the rationales by the elements of the ground set – which is employed in Definition 3 only to show the similarities between the CSMR approach and the RMR model – bears no significance, because the DM can maximize *any* of the available linear orders in $\mathscr{L}$ to explain selection from a menu. The missing connection between menus and rationales is mentioned by the authors (Kalai et al., 2002, p. 2287): "*We fully acknowledge the crudeness of our approach. [...] More research is needed to define and investigate 'structured' forms of rationalization.*" Following the authors' comments, we can refer to an RMR as a *context-free multi-rationalization*.

On the contrary, in a CSMR there is a structural connection between menus and rationalizing linear orders, which takes into account both attractiveness and informativeness of items. Specifically, Definition 2 discloses the relationship between informativeness of alternatives and DM's preferences by requiring that the rationale justifying the selection from a menu A can only be triggered by an item belonging to A. That is why we call our approach 'context-sensitive'.

It is worth observing that the type of connection embodied by Definition 2 has been reported in several empirical and experimental studies. For instance, in marketing Huber et al. (1982) document that the appearance on the shelves of a product (the *decoy good*), whose attributes are all worse than those of a previously unchosen alternative, may lead the consumer to select the dominant alternative. Furthermore, Simonson (1989) and Kivetz et al. (2004) study the *compromise effect*, showing that the presence of products with extreme features (price, quality, etc.) induces the DM to select intermediate options.

On the other hand, Definition 2 gives no clue about why some item x in a menu A should catch the DM's attention, and guide her judgement in choosing an item from A by maximizing the rationale $\rhd_x$. This lack of additional conditions yields the non-testability of our context-sensitive paradigm:

Lemma 2. *Any choice has a context-sensitive multi-rationalization.*

2.2. An index of rationality

By Lemma 2, the general framework of our context-sensitive approach explains any observed choice. However, as the RMR model, it can be used to classify all types of choice behavior according to their 'level of rationality'. This is measured by the minimum number of linear orders that are needed in a CSMR of a choice: the larger this number, the less rational the behavior.

Definition 4. The *context-sensitive rationality index* $r_{\mathsf{ctx}}(c)$ of a choice c is the least number of distinct linear orders in a CSMR of c.

Clearly, $1 \leqslant r_{\mathsf{ctx}}(c) \leqslant |X|$ for any choice on X. In particular, c is rationalizable (in the classical sense) if and only if $r_{\mathsf{ctx}}(c) = 1$. At the other end of the spectrum of rationality lie all choices c such that $r_{\mathsf{ctx}}(c) = |X|$, which ought to be regarded as representative of a hardly justifiable – in fact, irrational – behavior.[6] For a ground set of small size, the computation of the context-sensitive rationality index is usually not very demanding.[7] The next example determines this rationality index for a choice on a set of size four.

Example 3. Let $c : \mathscr{X} \to X$ be the choice on $X = \{w, x, y, z\}$ defined by

$$wx\underline{y}z\,,\quad w\underline{x}y\,,\quad w\underline{x}z\,,\quad w\underline{y}z\,,\quad x\underline{y}z\,,\quad \underline{w}x\,,\quad w\underline{y}\,,\quad w\underline{z}\,,\quad \underline{x}y\,,\quad \underline{x}z\,,\quad \underline{y}z\,.$$

We show that $r_{\mathsf{ctx}}(c) = 2$. Axiom α fails for c, hence $2 \leqslant r_{\mathsf{ctx}}(c) \leqslant 4$. Moreover, a CSMR of c with two distinct rationales is given by the family $\mathscr{L} = \{\rhd_w, \rhd_x, \rhd_y, \rhd_z\}$ such that

$$\rhd_w = \rhd_y\,,\quad \rhd_x = \rhd_z\,,\quad y \rhd_w w \rhd_w x \rhd_w z\,,\quad x \rhd_x y \rhd_x z \rhd_x w\,.$$

Our rationality index is the context-sensitive counterpart of the one defined for RMR:

Definition 5 (*Kalai et al., 2002*)**.** The *context-free rationality index* $r(c)$ of a choice c is the least number of distinct linear orders in an RMR of c.[8]

Due to the non-testability of the RMR model, Kalai et al. (2002) anchor their analysis on this rationality index, and prove two facts (as usual, $|X| = n \geqslant 2$):

Proposition 1. $1 \leqslant r(c) \leqslant n-1$ *for any choice* c *on* X.

Proposition 2. *As* n *goes to infinity, the fraction of choices on* X *with* $r(c) = n-1$ *tends to* 1.

In words, there is a partition of all potential choices on a ground set with n elements into exactly $n-1$ classes of rationality, and the most irrational class (requiring $n-1$ rationales) collects all choices as n diverges. Proposition 2 has a nontrivial proof, but it is hardly surprising: in fact, it is expected that extremely irrational choices become the norm as the size of the ground set becomes overwhelmingly large.

However, one would also expect that the types of pathological behavior collected in the most irrational class start appearing only when X is sufficiently large. Regrettably, this is not the case. For instance, Luce and Raiffa's dinner, which encodes a choice on 3 items, does belong to the most irrational class, and yet it carries a natural justification. As we discuss later on, a similar problem arises for several other choices on 3 or 4 elements, which, despite belonging to the most irrational class, bear intuitive explanations. This evidence shows that a context-free approach may be unable to provide a fine distinction among levels of choice rationality.

A question arises: *Does a context-sensitive multi-rationalization provide a fine classification of all potential choices, in a way that the class requiring the maximum number of rationales is really expressive of a pathological behavior?* The next section addresses this issue: in fact, we shall show that the most irrational choice behavior is quite rare for small sets of alternatives, but becomes the norm for large ground sets.

[5] This is not the original definition given by Kalai et al. (2002). In fact, they simply say that there is a family of linear orders that can be used to explain selection; in particular, no indexing of the rationales by the elements of the ground set X is employed. However, their Proposition 1 – in which the authors prove that $n-1$ linear orders suffice to explain choice behavior on a ground set of size n – ensures that the original notion of an RMR and Definition 3 are fully equivalent.

[6] Our index of rationality does not explicitly take into account the *distance* of rationales from the observed choice, intended as the fraction of selections that each linear order justifies. This way to measure deviations from rationality has already been considered in the literature: see, e.g., the *swaps index* proposed by Apesteguia and Ballester (2017).

[7] A low computational complexity is a desirable property of rationality measures, as discussed in Apesteguia and Ballester (2017).

[8] Again, this is not the terminology used by Kalai et al. (2002). However, it is justified by the fact that the authors call the RMR approach 'context-free'.

2.3. Moodiness

Definition 6. A choice $c : \mathscr{X} \to X$ is *moody* if $r_{\mathsf{ctx}}(c) = n$.

Moody choices are pathological: they describe DMs who justify their behavior by 'local' explanations for each occasion. In any context-sensitive rationalization of a moody dataset, each item of the ground set conveys a distinct indication about the DM's decision, persuading her to apply in some menus a preference different from those suggested by other alternatives.[9]

An inspection of choices on small ground sets shows that the inequality $r_{\mathsf{ctx}}(c) \leqslant n-1$ always holds. Thus one may wonder whether moody behavior be potentially realizable. The simple proof of Proposition 1 in Kalai et al. (2002) does not carry over our approach, because only preferences indexed by items in a menu can be adopted to explain selection. However, we do have:

Theorem 1 (Moodiness Exists). *There are moody choices on sufficiently large ground sets.*

The technical proof of Theorem 1 is presented in the Online Appendix. It uses the notion of a *flipped choice*, which is defined on a linearly ordered set X, and is such that selection from a menu systematically 'oscillates' from the best to the worst item. We show that any flipped choice on 39 elements always requires exactly 39 distinct rationales.[10]

Given the difficulty of finding a moody choice, Theorem 1 raises a new query about the ubiquity of moodiness when the size of the ground set grows larger and larger. Nevertheless – similarly to what Proposition 2 in Kalai et al. (2002) states for the RMR approach – we still have:

Theorem 2 (Moodiness Asymptotically Prevails). *The fraction of moody choices tends to one as the number of items in the ground set goes to infinity.*

We employ *Ramsey Theory* and other technical results to show that Theorem 2 holds true. Therefore, we present its long proof in the Online Appendix, where it is derived as a consequence of a more general result (Theorem 5). The statement of Theorem 2 is compatible with empirical evidence: consumers that face large assortments tend to show preference uncertainty and confusion, especially if they are unfamiliar with the product category; moreover, the wider the variety is, the more likely this irrationality surfaces (Chernev, 2003; Mogilner et al., 2008; Chernev et al., 2015).

An analogous conclusion cannot be drawn for the partition generated by a context-free approach. Indeed, *all* choices on $n = 3$ items can be explained by many well-known models of bounded rationality, such as *choice with limited attention* (Masatlioglu et al., 2012), *categorize-then-choose* (Manzini and Mariotti, 2012), *basic rationalization theory* (Cherepanov et al., 2014), and *overwhelming choice* (Lleras et al., 2017). Nevertheless, many of these choices do belong to the maximally irrational class identified by the RMR approach, which requires $n-1 = 2$ rationales for a context-free explanation. The situation is not very different on a ground set of size $n = 4$. Here the fraction of choices satisfying any of the models mentioned above is between $\frac{1}{3}$ and $\frac{3}{8}$ (Giarlotta et al., 2022a, Lemma 8), and yet many of these choices require the maximum number $n-1=3$ of context-free rationales.

On the contrary, the very involved proofs of Theorems 1, 2, and 5 strongly suggest:

Conjecture. *No choices that can be explained by some model of bounded rationality present in the literature are moody.*

[9] Note that a different notion of moody choice already exists: see Manzini and Mariotti (2010).

[10] Such a pathology may manifest for smaller ground sets, but we are unaware of that.

3. Choice by salience

Here we describe a specification of the context-sensitive paradigm of multi-rationalization, in which we investigate the behavioral mechanism that drives certain items to influence the DM's evaluation. Specifically, we assume that alternatives are ordered in the DM's perception according to a binary relation of salience, and the most salient items in the menu induce her to apply a preference in the selection process.

The effects of salient alternatives in consumer choices are studied by Bordalo et al. (2012, 2013). In their framework, salience is an increasing function of the distance of attributes (payoffs, prices, quality, etc.) from the average, and it distorts the DM's evaluation by inflating the relative utility weights attached to salient attributes of goods.[11] Instead, we encode salience by a binary relation describing the extent to which noticeable items draw the DM's attention. Thus, differently from Bordalo et al. (2013), we define a notion of salience for items, rather than for attributes; moreover, we only give an ordinal priority of consideration, rather than a cardinal evaluation of salience.

3.1. The model

Definition 7. A *rationalization by salience* of a choice $c : \mathscr{X} \to X$ is a pair $\langle \mathscr{L}, \succsim \rangle$, where

(S1) $\mathscr{L} = \{\rhd_x : x \in X\}$ is a family of linear orders on X (the *rationales*),
(S2) $\succsim$ is a weak order on X (the *salience order*), and
(S3) $\rhd_x$ equals $\rhd_y$ whenever $x \sim y$ (the *normality condition*),

such that, for any $A \in \mathscr{X}$, $c(A) = \max\left(A, \rhd_x\right)$ for some $x \in \max(A, \succsim)$. Henceforth, we refer to a choice that has a rationalization by salience as a *choice by salience (CS)*.

The process of rationalization by salience goes as follows. By condition S1, each alternative x in the ground set is attached a linear order $\rhd_x$, which encodes the informativeness of x by ranking all items in X according to their attractiveness. To explain choice on a menu A, first the DM's attention is captured by the most salient items in it, which are the alternatives belonging to $\max(A, \succsim)$. Note that, by condition S2, the transitive and complete order $\succsim$ yields a partition of the ground set into equivalence classes of salience, which are linearly ordered by their degree of notability. Thus, the normality condition S3 ensures that the rationales associated to all items in $\max(A, \succsim)$ are the same. Finally, the DM selects an item by maximizing the rationale suggested by the maximally salient items in A.

The main assumption that makes the model of choice by salience testable is the normality condition S3, which says that equally salient items suggest identical criteria to apply in the selection process. From a theoretical point of view, this is a *uniformity assumption* about the impact of informativeness across all items in each class: a single rationale, which ranks alternatives according to their attractiveness, is attached to each equivalence class induced by the salience order.

There are several equivalent formalizations of a rationalization by salience as a two-stage procedure. These formalizations, which shed a different light on the model, appeal to either a suitable choice correspondence (called a *focusing filter*) or a suitable choice function (called a *focusing rule*). The next statement summarizes these two alternative formulations.

Lemma 3. *The following statements are equivalent for any choice $c : \mathscr{X} \to X$:*

(i) *c has a rationalization by salience;*

[11] In a similar direction goes another recent model based on salience (Lanzani, 2022).

(ii) *there are a family* $\mathscr{L} = \{\rhd_x : x \in X\}$ *of linear orders on* X *and a choice correspondence* $\Phi : \mathscr{X} \to \mathscr{X}$ *satisfying WARP such that* $c(A) = \max\left(A, \rhd_x\right)$ *for some* $x \in \Phi(A)$;[12]

(iii) *there are a family* $\mathscr{L} = \{\rhd_x : x \in X\}$ *of linear orders on* X *and a choice function* $d : \mathscr{X} \to X$ *satisfying Axiom* α *such that* $c(A) = \max\left(A, \rhd_{d(A)}\right)$ *for any* $A \in \mathscr{X}$.

Remark 1. As for many other bounded rationality models (Manzini and Mariotti, 2012; Masatlioglu et al., 2012; Cherepanov et al., 2014; Lleras et al., 2017), also a rationalization by salience explains choice behavior by means of a two-stage procedure: first the DM focuses on the most salient alternatives, and successively she selects the item maximizing the preference suggested by them. However, differently from the above mentioned behavioral patterns, the first stage of a rationalization by salience does not necessarily shrink the set of choosable items: indeed, the DM may well choose an alternative that is not among the most salient ones. For instance, consider Luce and Raiffa's dinner (Examples 1 and 2), where the most salient item (frog's legs) is never selected from any nontrivial menu.

Remark 2. Choice by salience turns out to be equivalent to the model discussed by Kibris et al. (2021), who propose a theory of 'reference point formation'. Furthermore, it can also be considered as a special case of the general model of Lim (2021), which applies to risk, time, and social preferences. Yet another equivalent formulation of choice by salience is presented by Ravid and Stevenson (2021) within the framework of a 'temptation' model. However, our approach and the ones mentioned above were formulated totally independently.[13]

Next we associate to any choice a relation that is revealed by minimal irrational features.

Definition 8. Let $c : \mathscr{X} \to X$ be a choice. Define the relation $\vDash$ of *revealed salience* by $x \vDash y$ if there is a menu A containing y such that $(A, A \cup x)$ is a switch, for all distinct $x, y \in X$.

We write $x \vDash A$ whenever $(A, A \cup x)$ is a switch, because the latter fact implies $x \vDash a$ for all $a \in A$. Essentially, $\vDash$ infers salience from observed data: if adding x to a menu A causes a switch, then x is revealed to be more salient than any item in A.[14] In the path to characterizing a choice by salience, the following fact is algorithmically crucial:[15]

Lemma 4. *For any choice, if revealed salience is asymmetric, then it is also acyclic.*

In words, the absence of revealed cycles of length two suffices to prove the absence of revealed cycles of any length. The converse of Lemma 4 fails to hold: salience retrieved from choice can be acyclic, and not asymmetric.

Example 4. Define $c : \mathscr{X} \to X$ on $X = \{x, y, z\}$ by $xy\underline{z}$, $\underline{x}y$, $\underline{x}z$, $\underline{y}z$.

3.2. Salient attention

We shall show that choice by salience is equivalent to a special case of the following well-known model of bounded rationality:

Definition 9 (*Masatlioglu et al., 2012*)**.** A choice $c : \mathscr{X} \to X$ is *with limited attention (CLA)* if $c(A) = \max(\Gamma(A), \rhd)$ for all $A \in \mathscr{X}$, with $\rhd$ linear order on X, and $\Gamma : \mathscr{X} \to \mathscr{X}$ choice correspondence on X (*attention filter*) such that for all $B \in \mathscr{X}$ and $x \in X$,[16]

$$x \notin \Gamma(B) \implies \Gamma(B) \setminus x = \Gamma(B \setminus x). \tag{1}$$

A meaningful specification of CLA is the following:

Definition 10. A choice $c : \mathscr{X} \to X$ is *with salient limited attention (CSLA)* if it is CLA, and the attention filter Γ satisfies the following stronger condition for all $B \in \mathscr{X}$ and $x \in X$:

$$x \notin \{\min(B, \rhd), \max(\Gamma(B), \rhd)\} \implies \Gamma(B) \setminus x = \Gamma(B \setminus x). \tag{2}$$

In this case, Γ is called a *salient attention filter*.

In the CSLA model, the DM selects a unique item from a menu A by maximizing a linear order $\rhd$ on the submenu $\Gamma(A)$ determined by a salient attention filter Γ.[17] The DM's consideration is only affected by items holding an extreme position in her judgement, either maximum or minimum. In fact, condition (2) in Definition 10 states that for any menu B, if $x \in B$ is different from the extreme items $\min(B, \rhd)$ and $\max(\Gamma_s(B), \rhd)$, then removing x from B does not cause any change in the salient attention filter.

The latter feature of a salient attention filter is coherent with the salience theory of choice under risk of Bordalo et al. (2012): the DM's evaluation of lotteries is affected by extreme payoffs, which makes her risk-loving when upsides are high, and a risk-averse if downsides are high. Similarly, in Bordalo et al. (2013), items whose quality or price are far from the average distort consumer's perception.

3.3. Characterization

Here we state two characterizations of choice by salience. The first characterization relies on the binary relation of revealed salience: its mere asymmetry is necessary and sufficient for having a choice by salience. The second characterization explicitly identifies choice by salience with the specification of choice with limited attention determined by a salient attention filter.[18]

Theorem 3 (Characterization of Choice by Salience)**.** *The following statements are equivalent for a choice* c*:*

(i) c *is a choice by salience;*
(ii) *revealed salience is asymmetric;*
(iii) c *is a choice with salient limited attention.*

[12] A choice correspondence $\Phi : \mathscr{X} \to \mathscr{X}$ satisfies WARP when for all $A, B \in \mathscr{X}$ and $x, y \in X$, if $x, y \in A \cap B$, $x \in \Phi(A)$, and $y \in \Phi(B)$, then $x \in \Phi(B)$. By the *Fundamental Theorem of Revealed Preference Theory* – see, e.g., Sen (1971) – a choice correspondence satisfies WARP if and only if it is rationalizable by a weak order.

[13] In addition to the fact that choice by salience is a specification of the (novel) paradigm of context-sensitive rationalization, note that a preliminary draft of this paper was presented by Angelo Petralia in 2019, during his visit at the Universitat Pompeu Fabra of Barcellona.

[14] Revealed salience is called *revealed conspicuity* in the reference-dependence theory of Kibris et al. (2021). In fact, $\vDash$ is the converse of the relation $\widetilde{P}$ defined in Ravid and Stevenson (2021). See also Dutta and Horan (2015).

[15] None of the models that are equivalent to choice by salience points out this fact.

[16] The condition for an attention filter given by Masatlioglu et al. (2012) is that $x \notin \Gamma(B)$ implies $\Gamma(B) = \Gamma(B \setminus x)$. We use instead the equivalent condition (1) to better compare it with condition (2), which defines the notion of *salient* attention filter.

[17] It is worth noting that Definition 10 makes explicit the dependence of the attention filter on the DM's rationale. This dependence is implicit in the CLA model, and becomes explicit when constructing an attention filter from the given rational: see the proof of Theorem 3 in Masatlioglu et al. (2012, p. 2202).

[18] It is also possible to characterize choice by salience by the satisfaction of a testable axiom of choice consistency, called *WARP(S) (WARP by Salience)*, similarly to the characterization of CLA by means of WARP(LA) given in Masatlioglu et al. (2012). However, we prefer to focus our attention on the two equivalences stated in Theorem 3.

Lemma 4 and Theorem 3 show that the parameter identification in a rationalization by salience is computationally easy, and this fact is crucial in applications. By comparison, parameter identification in less restrictive models – such as those of Masatlioglu et al. (2012), and Cherepanov et al. (2014) – is more demanding. While the characterization of these theories effectively amounts to the requirement that the associated revealed preferences do not contain cycles of *arbitrary* length, choice by salience is simply characterized by the absence of cycles of length two. Moreover, this requirement is less involved than those offered by Kibris et al. (2021) and Ravid and Stevenson (2021), respectively based on the so-called *Single Reversal Axiom* and the *Axiom of Revealed Temptation*.

A CSLA representation of a choice by salience also offers an alternative interpretation of choice data. In fact, we have:

Lemma 5. *Let $c : \mathscr{X} \to X$ be a choice with salient limited attention. There is an explanation $(\Gamma, \rhd)$ of c such that, if there are $A \in \mathscr{X}$ and $x, y \in A$ for which $y \neq c(A) \neq c(A \setminus y)$, then $x \rhd y$ and $y = \min(\Gamma(A), \rhd)$.*

In other words, for a CSLA, if removing the item y from a menu A containing the item x causes a switch, then we can deduce the following facts: (i) the DM prefers x to y, and pays attention to y at A; (ii) y is the least preferred item among those brought to her attention in A. Thus, there are at most two alternatives that matter to the DM: the selected item, and the least preferred one (among those observed). In terms of tractability, this feature allows the analyst to identify any menu by two representative elements.

3.4. Numerical estimates

Here we show that choice by salience is a selective bounded rationality model, even when the number of items in the ground set is rather small. To that end, we start by observing that choices by salience are never moody.[19] By Theorem 2, the fraction of choices by salience goes to zero as the number of items diverges.

To offer a more concrete analysis, next we evaluate the fraction of choices by salience for some specific sizes of the ground set. All estimates are obtained by using the methodology introduced in Giarlotta et al. (2022a) and technically analyzed in Giarlotta et al. (2022b): we refer the reader to those two papers for details.

Definition 11. A *subchoice* of a choice $c : \mathscr{X} \to X$ is any choice $c_{\upharpoonright A} : \mathscr{A} \to A$, with $A \in \mathscr{X}$ and $\mathscr{A} := \{B \in \mathscr{X} : B \subseteq A\}$, defined by $c_{\upharpoonright A}(B) = c(B)$ for all $B \in \mathscr{A}$.

Definition 12. Two choices $c : \mathscr{X} \to X$ and $c' : \mathscr{X}' \to X'$ are *isomorphic* if there is a bijection $\sigma : X \to X'$ (called an *isomorphism*) such that $\sigma(c(A)) = c'(\sigma(A))$ for any $A \in \mathscr{X}$.

Definition 13. A *property $\mathscr{P}$ of choices* is a set of choices closed under isomorphism. We denote by $T(n)$, $T(n, \mathscr{P})$, and $F(n, \mathscr{P}) = \frac{T(n,\mathscr{P})}{T(n)}$, respectively, the total number of choices on n elements, the total number of choices on n elements satisfying property $\mathscr{P}$, and the fraction of choices on n elements satisfying property $\mathscr{P}$.

The ratio $F(n, \mathscr{P}) = \frac{T(n,\mathscr{P})}{T(n)}$ can be computed only considering choices on n elements that are pairwise non-isomorphic, because all isomorphism classes have exactly the same size (which is equal to $n!$): see Giarlotta et al. (2022a, Lemma 4).

Definition 14. A property $\mathscr{P}$ of choices is *hereditary* whenever if $\mathscr{P}$ holds for any choice, then it also holds for any of its subchoices.[20]

[19] The proof of this simple fact is left to the reader.

[20] Thus, $\mathscr{P}$ is hereditary if for all choices $c : \mathscr{X} \to X$, $c \in \mathscr{P}$ implies $c_{\upharpoonright A} \in \mathscr{P}$ for all $A \in \mathscr{X}$.

Lemma 6 (*Giarlotta et al. (2022a), Corollary 5*). *If $\mathscr{P}$ is a hereditary property that contains at most q pairwise non-isomorphic choices on four elements, then the following upper bounds to $F(n, \mathscr{P})$ hold:*

n	4	16	20	28	32
$F(n, \mathscr{P})$	$=(q/864)$	$\leqslant(q/864)^{20}$	$\leqslant(q/864)^{29}$	$\leqslant(q/864)^{57}$	$\leqslant(q/864)^{72}$

It is not difficult to show that the following facts hold:

Lemma 7. *The class of choices by salience is hereditary. Moreover, there are exactly 40 pairwise non-isomorphic choices by salience on four elements.*

In comparison, there are exactly 864 pairwise non-isomorphic choices on four items, of which 324 are CLA (Giarlotta et al., 2022a, Lemma 8), and only 1 is rationalizable. Lemmata 6 and 7 readily yield the numerical estimates we were after, which explicitly show the selectivity of a rationalization by salience:

Theorem 4 (Selectivity of Choice by Salience). *The following upper bounds hold for the fractions $F(n, \mathscr{P})$ of choices on $n = 4, 16, 20, 28$ elements, which are rationalizable, choices by saliences, or choices with limited attention:*

$\mathscr{P}$ \ n	4	16	20	28	32
Axiom α	$= 0.0011$	$\leqslant 10^{-58}$	$\leqslant 10^{-85}$	$\leqslant 10^{-167}$	$\leqslant 10^{-211}$
CS	$= 0.046$	$\leqslant 10^{-26}$	$\leqslant 10^{-38}$	$\leqslant 10^{-76}$	$\leqslant 10^{-96}$
CLA	$= 0.37$	$\leqslant 10^{-8}$	$\leqslant 10^{-12}$	$\leqslant 10^{-24}$	$\leqslant 10^{-30}$

Theorem 4 shows the model of choice by salience is much less restrictive than being retrievable by a binary relation; however, our model only explains a small portion of all choices with limited attention. Note that $F(n, \alpha)$, $F(n, \text{CS})$, and $F(n, \text{CLA})$ measure the relative fraction of all *potential* choices that are explained, respectively, by rationalizability, salience, and limited attention.[21] On the other hand, to assess the empirical validity of these models, one should consider the fraction of *actual* choices consistent with them.

4. Concluding remarks

In this paper we have provided a novel framework for context-sensitive choice behavior, in which the informativeness of some items may affect the individual perception of attractiveness of other (possibly different) items. The general version of our context-sensitive paradigm refines the RMR approach of Kalai et al. (2002) by providing a structured multi-rational explanation of choice behavior. The classification of choice functions into rationality levels prompted by our approach is empirically significant. Moreover, bounded rationality models arise from the context-sensitive paradigm as soon as we impose sound constraints. In fact, we single out a testable model of salience, which identifies the subclass of choices with limited attention of Masatlioglu et al. (2012) such that only salient items affect the DM's consideration.

The analysis of this paper hinges on a deterministic representation of salience, which implies that the perceived salience of items remains constant across menus. Possible extensions should consider a stochastic approach to salience, attained by considering a probability distribution over possible binary representations of it. Moreover, although the assumption on the salience ordering is rather consolidated in the literature, the composition of menus may affect the role of the items in DM's perception, creating cycles of any length. Thus, another possible direction of research is to design weaker properties of salience, which consider reversals of salience caused by different combinations of alternatives in distinct menus.

[21] Selten (1991) calls this measure the *area of a theory*.

Declaration of competing interest

Alfio Giarlotta gratefully acknowledges financial support by the Research Grant PRIN 2017 (CUP: J44I20000250001) "Multiple Criteria Decision Analysis and Multiple Criteria Decision Theory".

Angelo Petralia acknowledges the support of "Ministero del Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), PE9 GRINS "Spoke 8", project *Growing, Resilient, INclusive, and Sustainable*, CUP E63C22002120006.

Data availability

No data was used for the research described in the article.

Appendix A. Proofs (except Theorems 1 and 2)

Proof of Lemma 1. Suppose there are $A, B \in \mathscr{X}$ such that (A, B) is a switch, hence $c(A) \neq c(B) \in A$. If $|B \setminus A| = 1$, the claim holds. Thus assume $|B \setminus A| > 1$. Then there are $x \in X$ and $C \in \mathscr{X}$ such that $A \subsetneq C \subsetneq C \cup x = B$. If (C, B) is a (minimal) switch, then we are done again. Thus suppose (C, B) is not a switch.

Claim: (A, C) *is a switch.* By hypothesis, either (i) $c(B) = x$ or (ii) $c(B) = c(C)$ holds. Since (A, B) is a switch, case (i) cannot happen, hence $c(B) = c(C) = b \neq x$. It follows that $b \in A \setminus c(A)$, because otherwise (A, B) would fail to be a switch, contradicting the hypothesis. This proves that (A, C) is a switch.

Now the original violation of Axiom α witnessed by the switch (A, B) takes place within the smaller pair (A, C), where $C = B \setminus x$. If (A, C) is minimal, then we are done. Otherwise, we can repeat the above argument and show that there are $y \in X$ and $D \in \mathscr{X}$ such that $A \subsetneq D \subsetneq D \cup y = C$, and either (A, D) or (D, C) is a switch. Since X is finite, we eventually obtain what we are after. □

Proof of Lemma 2. Let $c : \mathscr{X} \to X$ be a choice. Define $|X|$-many linear orders $\rhd_x$, with $x \in X$, with the property that x is the top element of $\rhd_x$. For any $A \in \mathscr{X}$, $c(A) = \max(A, \rhd_x)$ for some $x \in A$. □

Proof of Lemma 3. Straightforward. □

Proof of Lemma 4. Let $c : \mathscr{X} \to X$ be a choice, and $\vDash$ the relation of revealed salience. We first prove three preliminary results,namely Lemmata 8, 9, and 10. □

Lemma 8. *Let $A \in \mathscr{X}$ and $x, y \in X$ be such that $x \neq y \in A$ and $x \nvDash y$.*

(i) *If $x \notin A$, then adding x to A does not switch the choice, except maybe to x.*
(ii) *If $x \in A \setminus c(A)$, then removing x from A does not affect the choice.*

Proof. By Definition 8, $x \vDash y$ means that there is $A \in \mathscr{X}$ such that $y \in A$ and $c(A) \neq c(A \cup x) \neq x$. Thus $x \nvDash y$ means that for any $A \in \mathscr{X}$ containing y, $c(A \cup x)$ is equal to either x or $c(A)$. Now both (i) and (ii) readily follow. □

Lemma 9. *Let $A, A', B \in \mathscr{X}$ and $x \in X$ be such that $A' \subseteq A$ and $x \in B$. If $A \nvDash x$, then $c(B \cup A') \in A' \cup c(B)$.*

Proof. Take $A, A', B \in \mathscr{X}$ and $x \in B$ such that Lemma 9 fails, where A' is a subset of A that is minimal for this failure.[22] Thus, $A \nvDash x$ and $c(B \cup A') \notin A' \cup c(B)$. If $A' = \{y\}$ for some $y \in X$, then $y \nvDash x$ and $c(B \cup y) \notin \{c(B), y\}$. However, this is impossible by Lemma 8. Next, consider the case $|A'| \geqslant 2$. Choose $y \in A'$, and set $A'' := A' - y \subseteq A$. By the minimality of A', we get $c(B \cup A'') \in A'' \cup c(B) \subseteq A' \cup c(B)$. It follows that $c(B \cup A'') \neq c(B \cup A') = c(B \cup A'' \cup y)$, which contradicts $y \nvDash x$. □

[22] Alternatively, A' is the smallest subset of A, such that, given $A', B \in \mathscr{X}$, and $x \in B$, Lemma 9 fails.

Lemma 10 (*Choice on Triples*). *Suppose $\vDash$ is asymmetric. For any distinct $x, y, z \in X$, if $x \vDash y$ and $x \nvDash z \nvDash y$, then $c(xyz) \neq y$.*

Proof. Let x, y, z be distinct elements of X satisfying the hypothesis. Since $x \vDash y$, there is $A \in \mathscr{X}$ such that $x, y \notin A$, $c(A \cup y) \neq c(A \cup xy) \neq x$. Thus, we have $x \vDash A \cup y$, which in turn implies $y \nvDash x$ and $A \nvDash x$ by the asymmetry of $\vDash$. Note also that $z \notin A$, since otherwise $x \vDash z$, contradicting the hypothesis. Now we can make the following deductions:

(i) if $c(A \cup yz) \neq z$, then $c(A \cup yz) = c(A \cup y)$ (since $z \nvDash y$);
(ii) if $c(A \cup xyz) \neq x$, then $c(A \cup xyz) = c(A \cup yz)$ (since $x \nvDash z$);
(iii) if $c(A \cup xyz) \neq y$, then $c(A \cup xyz) = c(A \cup xz)$ (since $y \nvDash x$);
(iv) if $c(A \cup xyz) \neq z$, then $c(A \cup xyz) = c(A \cup xy)$ (since $z \nvDash y$);
(v) if $c(A \cup xyz) \notin A$, then $c(A \cup xyz) = c(xyz)$ (by Lemma 9, since $A \nvDash x$).

Three cases: (1) $c(A \cup xyz) \in A$; (2) $c(A \cup xyz) = y$; (3) $c(A \cup xyz) \in \{x, z\}$.

In case (1), the implications (ii), (iii), and (iv) yield $c(A \cup yz) = c(A \cup xz) = c(A \cup xy)$, hence these chosen items are all equal to some $a \in A$. Now (i) applies, and so $c(A \cup y) = a$, which contradicts $c(A \cup xy) \neq c(A \cup y)$. In case (2), the implications (ii), (iv), and (v) yield $c(A \cup yz) = c(A \cup xy) = c(xyz)$, hence these chosen items are all equal to y. Now (i) applies, and so $c(A \cup y) = y$, which again contradicts $c(A \cup xy) \neq c(A \cup y)$. It follows that case (3) holds, and so the implication (v) yields $c(A \cup xyz) = c(xyz)$. This entails $c(xyz) \neq y$, thus completing the proof of Lemma 10. □

We now prove Lemma 4. Toward a contradiction, suppose $\vDash$ is asymmetric, but there is a $\vDash$-cycle of minimum length, say $a_1 \vDash a_2 \vDash \ldots \vDash a_n \vDash a_1$, where $n \geqslant 3$ and all a_i's are distinct; let $C = \{a_1, a_2, \ldots, a_n\}$ be the set of items involved in the cycle. To start, assume $n = 3$, that is, $a_1 \vDash a_2 \vDash a_3 \vDash a_1$ and $C = \{a_1, a_2, a_3\}$. By the asymmetry of $\vDash$ and Lemma 10, we get:

(1) $a_1 \vDash a_2$ and $a_1 \nvDash a_3 \nvDash a_2$, hence $c(a_1a_2a_3) \neq a_2$;
(2) $a_2 \vDash a_3$ and $a_2 \nvDash a_1 \nvDash a_3$, hence $c(a_1a_2a_3) \neq a_3$;
(3) $a_3 \vDash a_1$ and $a_3 \nvDash a_2 \nvDash a_1$, hence $c(a_1a_2a_3) \neq a_1$.

Thus $c(C)$ is empty, a contradiction.

Next, assume $n = 4$, i.e., $a_1 \vDash a_2 \vDash a_3 \vDash a_4 \vDash a_1$ and $C = \{a_1, a_2, a_3, a_4\}$. Minimality yields $a_1 \nvDash a_3 \nvDash a_1$ and $a_2 \nvDash a_4 \nvDash a_2$, and asymmetry entails $a_1 \nvDash a_4 \nvDash a_3 \nvDash a_2 \nvDash a_1$. By Lemma 10, we make the following deductions:

(1) $a_1 \vDash a_2$ and $a_1 \nvDash a_3 \nvDash a_2$, hence $c(a_1a_2a_3) \neq a_2$;
(2) $a_2 \vDash a_3$ and $a_2 \nvDash a_1 \nvDash a_3$, hence $c(a_1a_2a_3) \neq a_3$;
(3) $a_2 \vDash a_3$ and $a_2 \nvDash a_4 \nvDash a_3$, hence $c(a_2a_3a_4) \neq a_3$;
(4) $a_3 \vDash a_4$ and $a_3 \nvDash a_2 \nvDash a_4$, hence $c(a_2a_3a_4) \neq a_4$;
(5) $a_3 \vDash a_4$ and $a_3 \nvDash a_1 \nvDash a_4$, hence $c(a_3a_4a_1) \neq a_4$;
(6) $a_4 \vDash a_1$ and $a_3 \nvDash a_2 \nvDash a_1$, hence $c(a_3a_4a_1) \neq a_1$.

Thus, we have $c(a_1a_2a_3) = a_1$, $c(a_2a_3a_4) = a_2$, and $c(a_3a_4a_1) = a_3$. We now derive that $c(C)$ is empty, a contradiction. Indeed, $a_4 \nvDash a_3$ implies that $c(C)$ is equal to either a_4 or $c(a_1a_2a_3) = a_1$. Similarly, $a_1 \nvDash a_4$ implies that $c(C)$ is equal to either a_1 or $c(a_2a_3a_4) = a_2$, and $a_2 \nvDash a_1$ implies that $c(C)$ is equal to either a_2 or $c(a_3a_4a_1) = a_3$. Summarizing, we get $c(C) \in \{a_1, a_2\} \cap \{a_1, a_3\} \cap \{a_2, a_3\} = \varnothing$, as claimed.

In the general case, let $a_1 \vDash a_2 \vDash \ldots \vDash a_n \vDash a_1$, with $C = \{a_1, a_2, \ldots, a_n\}$. By a similar argument, we get $c(a_1a_2 \ldots a_{n-1}) = a_1$, $c(a_2a_3 \ldots a_n) = a_2$, and $c(a_3a_4 \ldots a_1) = a_3$. Now, $a_n \nvDash a_{n-1}$ implies $c(C) \in \{a_1, a_n\}$, $a_1 \nvDash a_n$ implies $c(C) \in \{a_1, a_2\}$, and $a_2 \nvDash a_1$ implies $c(C) \in \{a_2, a_3\}$, and so $c(C) = \{a_1, a_{n-1}\} \cap \{a_1, a_2\} \cap \{a_2, a_3\} = \varnothing$, which is impossible. This completes the proof of Lemma 4. □

Proof of Theorem 3. Le $c : \mathscr{X} \to X$ be a choice, and $\vDash$ the relation of revealed salience. We prove four implications: (i) $\Longrightarrow$ (ii), (ii) $\Longrightarrow$ (i), (i) $\Longrightarrow$ (iii), and (iii) $\Longrightarrow$ (ii).[23]

(i) $\Longrightarrow$ (ii): We assume c has a rationalization by salience $\langle \mathscr{L}, \succsim \rangle$, and show that the relation $\vDash$ of revealed salience is asymmetric. $\square$

Lemma 11. *For any $A \in \mathscr{X}$ and $x \in X$, if $(A, A \cup x)$ is a switch, then $x \succ a$ for all $a \in A$.*

Proof. Toward a contradiction, suppose $w \succsim x$ for some $w \in A$, hence $\max(A, \succsim) \subseteq \max(A \cup x, \succsim)$. By normality, we get $c(A \cup x) = c(A)$ or $c(A \cup x) = x$, which is false. $\square$

By Lemma 11, $\succ$ is an asymmetric extension of $\vDash$. Thus $\vDash$ is asymmetric, too.

(ii) $\Longrightarrow$ (i): We need some preliminary results, namely Lemmata 12, 13, and 14.

Lemma 12. *If $\vDash$ is asymmetric, then its transitive closure can be extended to a weak order.*

Proof. Asymmetry of $\vDash$ implies its acyclicity by Lemma 4. By Szpilrajn (1930)'s theorem, there is a weak order extending the transitive closure of $\vDash$. $\square$

In what follows, $\succsim$ denotes a weak order that extends the transitive closure of $\vDash$, whereas $\succ$ is the strict part of $\succsim$. Furthermore, for any $x \in X$, set $x^{\downarrow} := \{y \in X : x \succsim y\}$.

Lemma 13. *If $\vDash$ is asymmetric, then any pair (A, B) of menus included in $x^{\downarrow}$ is not a switch, as long as x belongs to both A and B.*

Proof. Suppose $\vDash$ is asymmetric. Toward a contradiction, assume there are $x \in X$ and $A, B \in \mathscr{X}$, with $A \subsetneq B \subseteq x^{\downarrow}$, and $x \in A$, such that (A, B) is a switch. By Lemma 1, there are $y \in X$ and $C \in \mathscr{X}$ such that $A \subseteq C \subsetneq C \cup y \subseteq B$ and $(C, C \cup y)$ is a switch. It follows that $y \vDash C \supseteq A$, and so $y \succ A$. We conclude that $y \succ x$, a contradiction. $\square$

For any $x \in X$ and distinct $y, z \in x^{\downarrow}$, define

$$y >_x z \quad \Longleftrightarrow \quad \text{there is } A \subseteq x^{\downarrow} \text{ such that } x \in A, z \in A \setminus c(A), \text{ and } y = c(A). \tag{3}$$

Note that if either y or z (or both) does not belong to $x^{\downarrow}$, then we leave y and z incomparable. Observe also that $>_x$ is irreflexive by construction. We abuse notation, and write $y >_x A$, whenever there is $A \subseteq x^{\downarrow}$ such that $x, y \in A$ and $y = c(A)$. The reason is that $y >_x A$ implies $y >_x z$ for any $z \in A \setminus y$.

Lemma 14. *If $\vDash$ is asymmetric, then $>_x$ is asymmetric and transitive for any $x \in X$.*

Proof. Assume $\vDash$ is asymmetric, and let $x \in X$. To prove that $>_x$ is asymmetric, suppose $y >_x z$ and $z >_x y$ for some $y, z \in X$. (Note that $y \neq z$, because $>_x$ is irreflexive.) By the definition of $>_x$, there are $A, B \subseteq x^{\downarrow}$ such that $x, y, z \in A \cap B$, $y = c(A)$, and $z = c(B)$. Consider the menu $A \cap B$, which is included in $x^{\downarrow}$ and contains x, y, z. If $c(A \cap B) \notin \{y, z\}$, then $(A \cap B, A)$ is a switch, which contradicts Lemma 13. On the other hand, if $c(A \cap B) = y$ (resp. $c(A \cap B) = z$), then $(A \cap B, B)$ (resp. $(A \cap B, A)$) is a switch, which is forbidden by Lemma 13. Thus $c(A \cap B) = \varnothing$, which is impossible.

To prove $>_x$ is transitive, let $w, y, z \in X$ be such that $w >_x y >_x z$. By the definition of $>_x$, there are $A, B \subseteq x^{\downarrow}$ such that $x, y \in A \cap B$, $z \in B$, $w = c(A)$, and $y = c(B)$. Consider the menu $A \cup B$, which is included in $x^{\downarrow}$ and contains x. We claim that $c(A \cup B) = w$. Indeed, if $c(A \cup B) \in (A \cup B) \setminus \{w, y\}$, then either $(A, A \cup B)$ or $(B, A \cup B)$ is a switch, which contradicts Lemma 13. Moreover, if $c(A \cup B) = y$, then $(A, A \cup B)$ is a switch, which is again impossible. This proves the claim. Now we get $w >_x (A \cup B)$, hence $w >_x z$. $\square$

We are ready to show that (ii) implies (i). Suppose $\vDash$ is asymmetric, and let $\succsim$ be a weak order extending the transitive closure of $\vDash$, which exists by Lemma 12. For any $x \in X$, define $>_x$ as in (3). By Lemma 14, each $>_x$ is asymmetric and transitive. Let $\rhd_x$ be a linear extension of $>_x$, which exists by Szpilrajn (1930)'s theorem. To complete the proof, we show that the pair $\langle \mathscr{L}, \succsim \rangle$, where $\mathscr{L} = \{\rhd_x : x \in X\}$, is a rationalization by salience of c. Let A be a menu, and x an item in $\max(A, \succsim)$. Note that $A \subseteq x^{\downarrow}$. By construction, $c(A) >_x A$, hence we can conclude $c(A) = \max\left(A, \rhd_x\right)$, as wanted.

(i) $\Longrightarrow$ (iii): We suppose $c : \mathscr{X} \to X$ has a rationalization by salience, and show that it is a choice with salient limited attention. Let $\widetilde{P}$ be the *converse* of $\vDash$, that is, $x \widetilde{P} y$ if and only if $y \vDash x$, for any $x, y \in X$. By Theorem 3 and Lemma 4, $\widetilde{P}$ is acyclic and asymmetric. Let $\rhd$ be any linear order that extends $\widetilde{P}$. Denoted $x^{\downarrow} := \{y \in X : x \rhd y \text{ or } y = x\}$ for any $x \in X$, define a choice correspondence $\Gamma_{\rhd} : \mathscr{X} \to \mathscr{X}$ as follows for all $A \in \mathscr{X}$:

$$\Gamma_{\rhd}(A) := c(A)^{\downarrow} \cap A. \tag{4}$$

Claim: (a) $c(A) = \max(\Gamma_{\rhd}(A), \rhd)$ *for all* $A \in \mathscr{X}$. (b) $\Gamma_{\rhd}$ *is a salient attention filter.*

This will show that c is CSLA. Part (a) readily follows from the definition of $\Gamma_{\rhd}$. To prove (b), let $B \in \mathscr{X}$ and $x \in X$. We deal separately with the two possible cases: (b1) $x \notin \Gamma_{\rhd}(B)$, and (b2) $x \in \Gamma_{\rhd}(B)$, but $x \neq \min(B, \rhd), \max(\Gamma_{\rhd}(B), \rhd)$.

Case b1: By (4), we get $x \rhd c(B)$. Since $\rhd$ extends $\widetilde{P}$, we derive that $x \vDash c(B)$ fails to hold, and so there is no menu $D \in \mathscr{X}$ such that $c(B) \in D$ and $x \neq c(D) \neq c(D \setminus x)$. It follows that $c(B) = c(B \setminus x)$, since otherwise $D := B$ would be a menu witnessing $x \vDash c(B)$, which is impossible. Now the definition of $\Gamma_{\rhd}$ and the hypothesis $x \notin \Gamma_{\rhd}(B)$ yield $\Gamma_{\rhd}(B) \setminus x = \Gamma_{\rhd}(B) = \Gamma_{\rhd}(B \setminus x)$, as claimed.

Case b2: Since $x \neq \max(\Gamma_{\rhd}(B), \rhd)$, formula (4) gives $c(B) \rhd x$. We claim that $c(B) = c(B \setminus x)$. Indeed, if $c(B) \neq c(B \setminus x)$, then $(B \setminus x, B)$ is a switch, hence $x \vDash (B \setminus x)$. It follows that $(B \setminus x) \rhd x$, and so $x = \min(B, \rhd)$, which contradict the hypothesis. By (4) and the claim, we get $\Gamma(B) \setminus x = (c(B)^{\downarrow} \cap B) \setminus x = c(B \setminus x)^{\downarrow} \cap (B \setminus x) = \Gamma(B \setminus x)$.

(iii) $\Longrightarrow$ (ii): Suppose $c : \mathscr{X} \to X$ is CSLA. Hereafter, we say that $(\Gamma, \rhd)$ *rationalizes* c if $c(A) = \max(\Gamma(A), \rhd)$ for all $A \in \mathscr{X}$, where $\rhd$ is a linear order on X, and Γ is a salient attention filter. In particular, $(\Gamma, \rhd)$ *maximally rationalizes* c if $(\Gamma, \rhd)$ rationalizes c, and there is no salient attention filter $\Gamma' : \mathscr{X} \to \mathscr{X}$ distinct from Γ such that $(\Gamma', \rhd)$ rationalizes c and $\Gamma(A) \subseteq \Gamma'(A)$ for all $A \in \mathscr{X}$.

Lemma 15. *If $(\Gamma, \rhd)$ rationalizes c, then $(\Gamma_{\rhd}, \rhd)$ maximally rationalizes c.*

Proof. Suppose $(\Gamma, \rhd)$ rationalizes c. We prove the following: (a) $c(A) = \max(\Gamma_{\rhd}(A), \rhd)$ for all $A \in \mathscr{X}$; (b) $\Gamma_{\rhd}$ is a salient attention filter; (c) $\Gamma_{\rhd}$ is maximal. Part (a) follows from the definition (4) of $\Gamma_{\rhd}$. For (b), let $B \in \mathscr{X}$ and $x \in B$ different from $\min(B, \rhd)$ and $\max(\Gamma_{\rhd}(B), \rhd)$. Toward a contradiction, suppose $\Gamma_{\rhd}(B) \setminus x \neq \Gamma_{\rhd}(B \setminus x)$. The definition of $\Gamma_{\rhd}$ yields $\left(c(B)^{\downarrow} \cap B\right) \setminus x \neq c(B \setminus x)^{\downarrow} \cap (B \setminus x)$, hence $c(B) \neq c(B \setminus x)$. Moreover, we have $x \neq \max(\Gamma(B), \rhd)$. Since $\Gamma(B) \setminus x = \Gamma(B \setminus x)$ because Γ is a salient attention filter, we obtain $c(B) \in \Gamma(B \setminus x)$ and $c(B \setminus x) \in \Gamma(B)$, which yield $c(B \setminus x) \rhd c(B)$ and $c(B) \rhd c(B \setminus x)$, a contradiction. To prove (c), suppose there is a salient attention filter Γ' such that $(\Gamma', \rhd)$ rationalizes c and $y \in \Gamma'(D) \setminus \Gamma_{\rhd}(D)$ for some $D \in \mathscr{X}$ and $y \in D$. Since $y \notin \Gamma_{\rhd}(D)$, we get $y \rhd c(D)$. On the other hand, since $y \in \Gamma'(D)$ and $(\Gamma', \rhd)$ rationalizes c, we must have $c(D) \rhd y$ or $c(D) = y$, which is impossible. $\square$

[23] We thank Davide Carpentiere for providing the proof of the last implication.

Lemma 16. *If* $(\Gamma_{\rhd}, \rhd)$ *maximally rationalizes* c*, then* $\rhd$ *extends* $\widetilde{P}$.

Proof. Suppose $(\Gamma_{\rhd}, \rhd)$ maximally rationalizes c. To show that $\rhd$ extends $\widetilde{P}$, we prove that $\neg(x \rhd y)$ implies $\neg(x \widetilde{P} y)$ for distinct $x, y \in X$. Suppose $\neg(x \rhd y)$, hence $y \rhd x$. Since $\widetilde{P}$ is the converse of $\vDash$, we need show that $y \nvDash x$. Toward a contradiction, suppose $y \vDash x$, that is, $y \neq c(B) \neq c(B \setminus y)$ for some menu $B \in \mathscr{X}$ containing x and y. Note that $y \neq \min(B, \rhd)$ and $y \neq \max(\Gamma_{\rhd}(B), \rhd) = c(B)$. Since c is CSLA, we obtain $\Gamma_{\rhd}(B) \setminus y = \Gamma_{\rhd}(B \setminus y)$, which implies that $c(B) \in \Gamma_{\rhd}(B \setminus y)$ and $c(B \setminus y) \in \Gamma_{\rhd}(B)$. Since $c(B) \neq c(B \setminus y)$, condition (4) yields $c(B \setminus y) \rhd c(B)$ and $c(B) \rhd c(B \setminus y)$, which is impossible. □

Now suppose $(\Gamma, \rhd)$ rationalizes c. By Lemmata 15 and 16, $\rhd$ extends $\widetilde{P}$. It follows that $\vDash$ is asymmetric, as claimed. This completes the proof of Theorem 3. □

Proof of Lemma 5. Straightforward. □

Proof of Lemma 7. This is similar to the proof of Lemma 8 in Giarlotta et al. (2022a), which is fully described in Giarlotta et al. (2022b), and can be easily adapted to the model of choice by salience. Details are available upon request. □

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmateco.2023.102913.

References

Apesteguia, J., Ballester, M.A., 2017. A measure of rationality and welfare. J. Polit. Econ. 123, 1278–1310.
Bhattachary, M., Mukherjee, S., Sonal, R., 2021. Frame-based stochastic choice rule. J. Math. Econom. 97, 102553.
Bordalo, P., Gennaioli, N., Shleifer, A., 2012. Salience theory of choice under risk. Q. J. Econ. 127, 1243–1285.
Bordalo, P., Gennaioli, N., Shleifer, A., 2013. Salience and consumer choice. J. Polit. Econ. 121, 803–843.
Cherepanov, V., Feddersen, T., Sandroni, A., 2014. Rationalization. Theor. Econ. 8, 775–800.
Chernev, A., 2003. The role of ideal point availability and assortment in consumer choice. J. Consum. Res. 30, 170–183.
Chernev, A., Böckenholt, U., Goodman, J., 2015. Choice overload: A conceptual review and meta-analysis. J. Consum. Psychol. 25, 333–358.
Chernoff, H., 1954. Rational selection of decision functions. Econometrica 22, 422–443.
Dutta, R., Horan, S., 2015. Inferring rationales from choice: Identification for rational shortlist methods. Am. Econ. J. Microecon. 7, 179–201.
Giarlotta, A., Petralia, A., Watson, S., 2022a. Bounded rationality is rare. J. Econom. Theory 204, 105509.
Giarlotta, A., Petralia, A., Watson, S., 2022b. On the number of non-isomorphic choices on four elements. MethodsX 9, 101854.
Goldin, J., Reck, D., 2020. Revealed-preference analysis with framing effects. J. Polit. Econ. 128, 2759–2795.
Huber, J., Payne, J.W., Puto, C., 1982. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. J. Consum. Res. 9, 90–98.
Kalai, G., Rubinstein, A., Spiegler, R., 2002. Rationalizing choice functions by multiple rationales. Econometrica 70, 2481–2488.
Kibris, O., Masatlioglu, Y., Suleymanov, E., 2021. A theory of reference point formation. Econom. Theory http://dx.doi.org/10.1007/s00199-021-01392-3.
Kivetz, R., Netzer, O., Srinivasan, V., 2004. Alternative models for capturing the compromise effect. J. Mar. Res. 41, 237–257.
Lanzani, G., 2022. Correlation made simple: Application to salience and regret theory. Q. J. Econ. 137, 959–987.
Lim, X.Z., 2021. Ordered reference dependent choice. arXiv:2105.12915 [econ.TH].
Lleras, J., Masatlioglu, Y.S., Nakajima, D., Ozbay, E.Y., 2017. When more is less: Limited consideration. J. Econom. Theory 170, 70–85.
Luce, D.R., Raiffa, H., 1957. Games and Decisions: Introduction and Critical Survey. Wiley, New York.
Manzini, P., Mariotti, M., 2010. Moody choices. Available at SSRN: https://ssrn.com/abstract=1623319 or http://dx.doi.org/10.2139/ssrn.1623319.
Manzini, P., Mariotti, M., 2012. Categorize then choose: Boundedly rational choice and welfare. J. Eur. Econom. Assoc. 10, 1141–1165.
Masatlioglu, Y., Nakajima, D., Ozbay, E.Y., 2012. Revealed attention. Amer. Econ. Rev. 102, 2183–2205.
Mogilner, C., Rudnick, T., Iyengar, S.S., 2008. How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. J. Consum. Res. 35, 202–215.
Ravid, D., Stevenson, K., 2021. Bad temptation. J. Math. Econom. 95, 102480.
Salant, Y., Rubinstein, A., 2006. Choice with frames. Rev. Econ. Stud. 75, 1287–1296.
Samuelson, A.P., 1938. A note on the pure theory of consumer's behaviour. Economica 17, 61–71.
Selten, R., 1991. Properties of a measure of predictive success. Math. Social Sci. 21, 153–167.
Sen, A.K., 1971. Choice functions and revealed preference. Rev. Econom. Stud. 38, 307–317.
Simonson, I., 1989. Choice based on reasons: The case of attraction and compromise effects. J. Consum. Res. 16, 158–174.
Szpilrajn, E., 1930. Sur l'extension de l'ordre partiel. Fund. Math. 16, 386–389.
Taylor, S.E., Fiske, S.T., 1978. Salience, attention, and attribution: Top of the head phenomena. Adv. Exp. Soc. Psychol. 11, 249–288.
Tversky, A., Kahneman, D., 1974. Judgement under uncertainty: Heuristics and biases. Science 185, 1124–1131.